



## Cloning Scheme for Generating pCMV-NS35

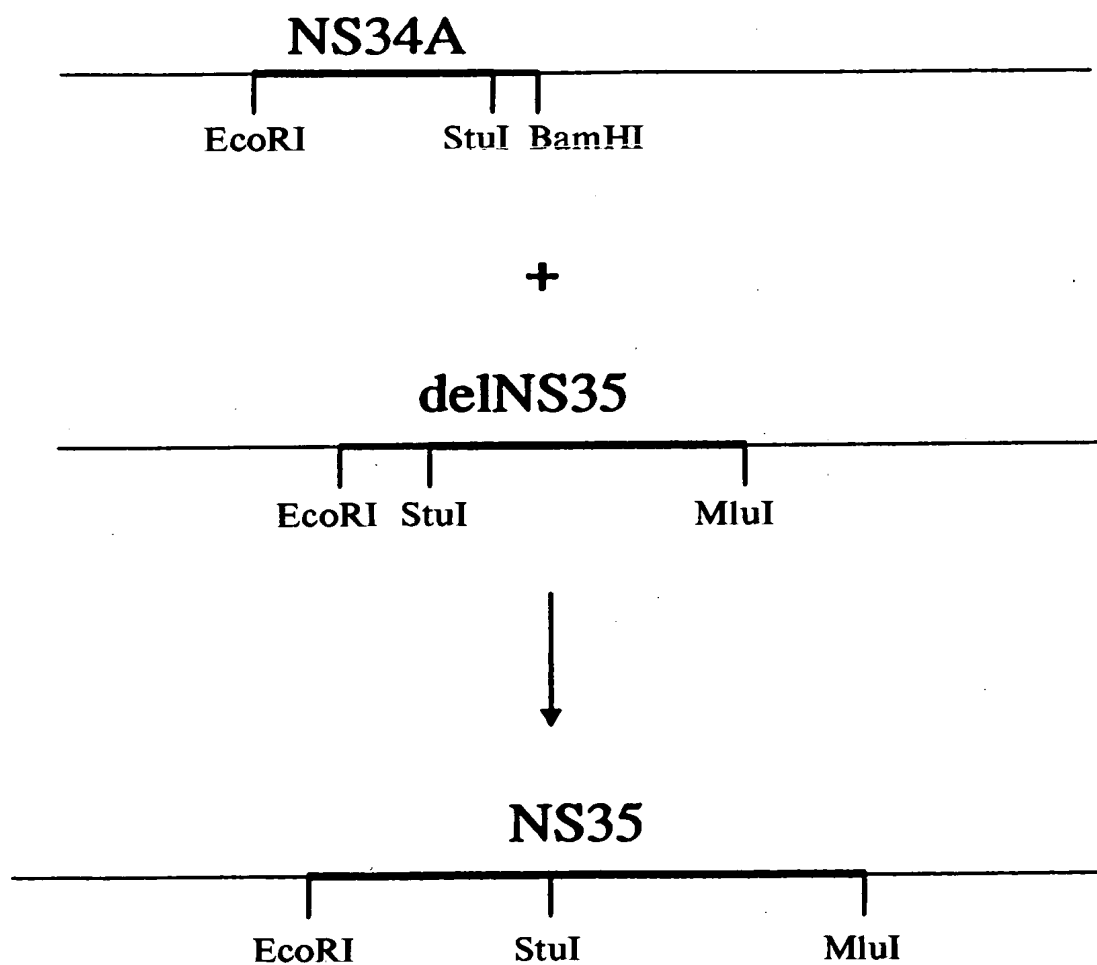


FIG. 1

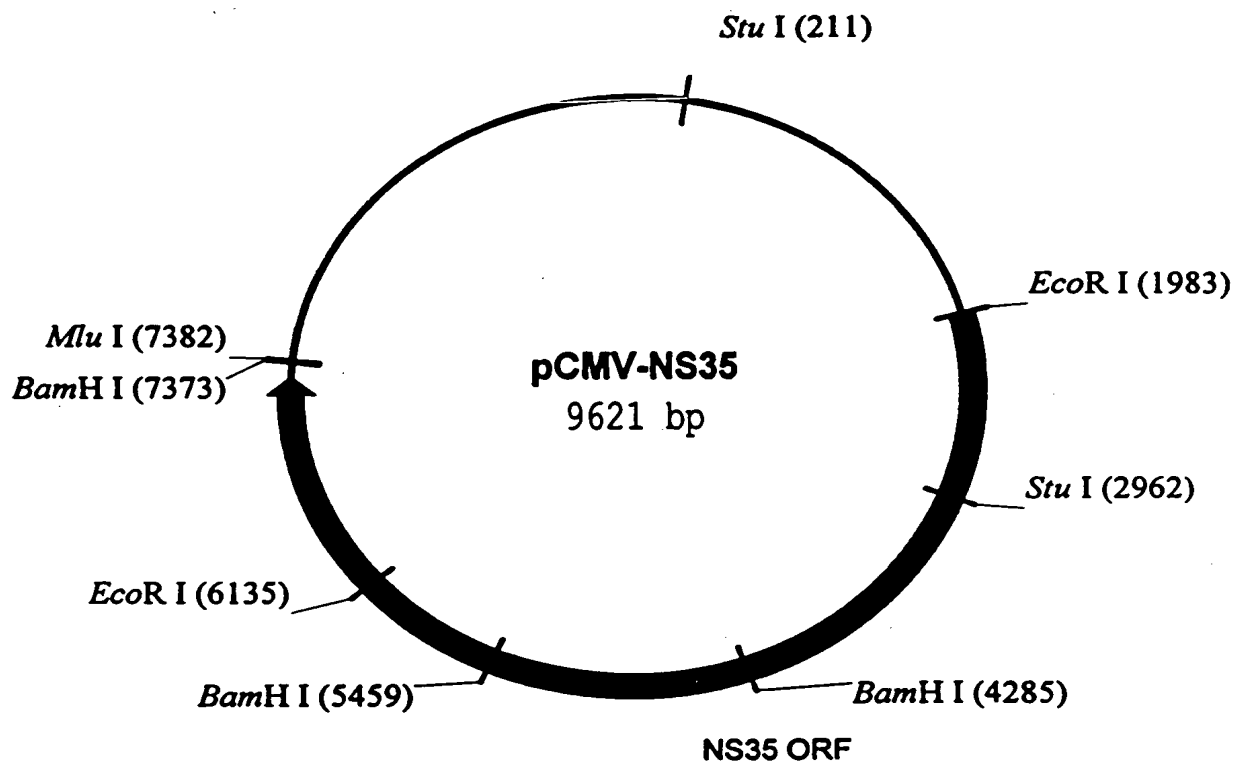


FIG. 2

PCMV-NS35

1	TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGTCT GTAAGCGGAT AGCGCGCAAA GCCACTACTG CCACCTTTGG AGACTGTGTA CGTCGAGGGC CTCTGCCAGT GTCGAACAGA CATTGCGCTA
81	GCGGGGAGCA GACAAGCCCG TCAGGGGGCG TCAGGGGGTG TTGGGGGGTG TCGGGGGCTGG CTTAACTATG CGGCATCAGA CGGCCCTCGT CTGTTCCGGC AGTCCCGCGC AGTCGCCAC AACCGCCAC AGCCCCGACC GAATTGATAC GCCGTAGTCT
161	GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTGGCA AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG CGTCTAACAT GACTCTCAG TGGTATACCT CGAAAAACGT TTTCCGGATCC GGAGGTTTT TCGGAGGAGT GATGAAGACC
241	AATAGCTCAG AGGCGGAGGC GGCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA TGGGGGGGAG AATGGCGGA TTATCGAGTC TCCGGTCCG CCGGAGCGG AGAGGTATTT ATTTTTTTA ATCAGTCGGT ACCCGCCTC TTACCCGCT
321	ACTGGGGGG GAGGGAATTA TTGGCTATTG GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT TGACCCGCCC CTCCCTTAAT AACCGATAAC CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA
401	CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTTCA GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA
481	AGCCCATATA TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCGGCCCAT TCGGGTATAT ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GGGCGGGTAA
561	GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTACGGT CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA
641	AAACTGCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG TCAATGACGG TAAATGGCCC TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGGCGG GGATAACTGC AGTTACTGCC ATTTACCGGG

StuI  
 -----

FIG. 3-1



pCMV-NS35

721 GCCTGGCATT ATGCCCAGTA CATGACCTTA CGGGACTTTC CTACTTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC  
CGGACCGTAA TACGGGTCAAT GTACTGGAAT GCCCTGAAAG GATGAACGGT CATGTAGATG CATAATCAGT AGCGATAATG

---

801 CATGGTGATG CGGTTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA  
GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATCGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGGTGGGT

---

881 TTGACGTCAA TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTGTA ATAACCCCGC CCCGTTGACG  
AACTGCAGTT ACCCTCAAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG GGGCAACTGC

---

961 CAAATGGCG GTAGCGGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG  
GTTTACCGCG CATCCGCACA TGCCACGCTC CAGATATATT CGTCTGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC

---

1041 CCATCCACGC TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGGCG GGAACGGTGC ATTGGAACGC  
GGTAGGTGG ACAAACCTGG AGGTATCTTC TGTGGCCCTG CTTAGGTCCG AGGGCGCGC CCTTGCCACG TAACCTTGCG

---

1121 GGATTCCCCG TGCCAAGAGT GACGTAAGTA CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA  
CCTAAGGGCG ACGGTTCTCA CTGCATTCTAT GCGGGATATC TGAGATATCC GTGTGGGGA ACCGAGAATA CGTACGATAT

---

1201 CTGTTTTTGG CTGGGGCCT ATACACCCCG GCTCCTTATG CTATAGGTGA TGGTATAGCT TAGCCTATAG GTGTGGGTTA  
GACAAAACC GAACCCCGGA TATGTGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCCAAT

---

1281 TTGACCATT TTAGCCACTC CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG CCACAACCTAT  
AACTGGTAAT AACTGGTGAG GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC GGTGTTGATA

---

1361 CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GACACGGACT CTGTATTTT ACAGGATGG GTCCATTAT  
GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCTCTGA GACATAAAAA TGTCCTACCC CAGGTAATA

FIG. 3-2

pCMV-NS35

```

1441 TATTTACAAA TTCACATATA CAACAACGCC GTCCCCCGTG CCGGCAGTTT TTATTAAACA TAGCGTGGGA TCTCCGACAT
    ATAAATGTTT AAGTGATAT GTTGTGCGG CAGGGGGCAC GGGCGTCAAA AATAATTGT ATCGCACCTT AGAGGCTGTA

1521 CTCGGGTACG TGTTCCGGAC ATGGGCTCTT CTCCGGTAGC GCGGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA
    GAGCCCATGC ACAAGGCCCTG TACCCGAGAA GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

1601 GCGGCTCATG GTCGCTCGGC AGTCTCTTGC TCCTAACAGT GGAGGGCAGA CTTAGGCACA GCACAATGCC CACCACCACC
    CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT GAATCCGTGT CGTGTIACGG GTGGTGGTGG

1681 AGTGTGCCG ACAAGGCCGT GCGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT GGACGCAGAT
    TCACACGGCG TGTTCGGCA CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGA CCTGCGTCTA

1761 GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT
    CCTTCTGAAT TCCGTGCGCG TCTTCTTCTA CGTCCGTGCA CTCACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA

1841 TCGGCTGCTG TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGGCGG CGCCACCAGA CATAATAGCT
    ACGCACGAC AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGGGCGG GCGGTGGTCT GTATTATCGA

+2                                     M A A
                                     EcoRI
                                     -----
1921 GACAGACTAA CAGACTGTTT CTTTCCATGG GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTACCC ATGGCTGCAT
    CTGTCTGATT GTCTGACAAG GAAAGGTACC CAGAAAAGAC CTCAGTGGA GCAGCTGGAT TCTTAAGTGG TACCGACGTA

+2 Y A A Q G Y K V L V L N P S V A A T L G F G A Y M S K
2001 ATGCAGCTCA GGGCTATAAG GTGCTAGTAC TCAACCCCTC TGTTGCTGCA ACACCTGGCT TTGGTGCTTA CATGTCCAAG
    TAGGTCGAGT CCCGATATTC CAGGATCATG AGTTGGGGAG ACAACGACGT TGTGACCCGA AACCCAGAAAT GTACAGGTTT

```

FIG. 3-3

pCMV-NS35

+2 A H G I D P N I R T G V R T I T T G S P I T Y S T Y G  
 2081 GCTCATGGGA TCGATCCTAA CATCAGGACC GGGGTGAGAA CAATTACCAC TGGCAGCCCC ATCAGGTACT CCACCTACGG  
 CGAGTACCCT AGCTAGGATT GTAGTCTGG CCCCACTCTT GTTAATGGTG ACCGTCGGG TAGTGCATGA GGTGGATGCC

+2 K F L A D G G C S G G A Y D I I I C D E C H S T D A  
 2161 CAAGTTCCTT GCCGACGGG GGTGCTCGG GGGGGCTTAT GACATAATAA TTTGTGACGA GTGCCACTCC ACGGATGCCA  
 GTTCAAGGAA CGGCTGCCG CCACGAGCCC CCCGGGAATA CTGTATTATT AAACACTGCT CACGGTGAGG TGCCTACGGT

+2 T S I L G I G T V L D Q A E T A G A R L V V L A T A T  
 2241 CATCCATCTT GGGCATTGGC ACTGTCTTG ACCAAGCAGA GACTGCGGG GCGAGACTGG TTGTGCTCGC CACCGCCACC  
 GTAGGTAGAA CCCGTAACCG TGACAGGAAC TGGTTCGTCT CTGACGCCCC CGCTCTGACC AACACGAGCG GTGGCGGTGG

+2 P P G S V T V P H P N I E E V A L S T T G E I P F Y G  
 2321 CCTCGGGCT CCGTCACTGT GCCCATCC AACATCGAG AGTTGTCT GTCCACCACC GGAGAGATCC CTTTTTACGG  
 GGAGGCCGA GGCAGTGACA CGGGTAGGG TTGTAGCTCC TCCAACGAGA CAGGTGTGG CCTCTCTAGG GAAAAATGCC

+2 K A I P L E V I K G G R H L I F C H S K K K C D E L  
 2401 CAAGGCTATC CCCCTCGAAG TAATCAAGG GGGGAGACAT CTCATCTTCT GTCATTCAAA GAAGAAGTGC GACGAACCTG  
 GTCCGATAG GGGGAGCTT ATTAGTTCC CCCCTCTGA GAGTAGAGA CAGTAAGTTT CTTCTTCACG CTGCTTGAGC

+2 A A K L V A L G I N A V A Y Y R G L D V S V I P T S G  
 2481 CCGCAAAGCT GGTGCGATTG GGCATCAATG CCGTGGCCTA CTACCGCGGT CTTGACGTGT CCGTCATCCC GACCAGCGGC  
 GGGTTTGA CCAGGTAAC CCGTAGTAC GGCACCGGAT GATGGGCCA GAACTGCACA GGCAGTAGGG CTGGTCGCCC

+2 D V V V A T D A L M T G Y T G D F D S V I D C N T C  
 2561 GATGTTGTCG TCGTGGCAAC CGATGCCCTC ATGACCGGCT ATACCGGCGA CTTGACTCG GTGATAGACT GCAATACGTG  
 CTACAACAGC AGCACCGTTG GCTACGGGAG TACTGGCGA TATGGCCGCT GAAGCTGACC CACTATCTGA CGTTATGCAC

FIG. 3-4



pCMV-NS35

```
+2  V T Q T V D F S L D P T F T I E T I T L P Q D A V S
2641 TGTACCCAG ACAGTCGATT TCAGCCCTTGA CCCTACCTTC ACCATTGAGA CAATCAGCT CCCCCAAGAT GCTGTCTCCC
    ACAGTGGGTC TGTACGCTAA AGTCGGAAG AGTCGGAAGT GGGATGGAAG TGGTAACCTCT GTTAGTGGGA GGGGGTTCTA CGACAGAGGG

+2  R T Q R R G R T G R G K P G I Y R F V A P G E R P S G
2721 GCACTCAACG TCGGGGCAGG ACTGGCAGGG GGAAGCCAGG CATCTACAGA TTTGTGGCAC CGGGGGAGCG CCCCTCCGGC
    CGTGAGTTGC AGCCCCGTCC TGACCGTCCC CCTTCGGTCC GTAGATGTCT AAACACCCGTG GCCCCCTCGC GGGGAGGCGC

+2  M F D S S V L C E C Y D A G C A W Y E L T P A E T T V
2801 ATGTTGACT CGTCCGTCCT CTGTGAGTGC TATGACGAG GCTGTGCTTG GTATGAGCTC ACGCCGCGCG AGACTACAGT
    TACAAGCTGA GCAGGCAGGA GAACTCAGG ATACTGCGTC CGACACGAAC CATACTCGAG TCGGGGCGCG TCTGATGTCA

+2  R L R A Y M N T P G L P V C Q D H L E F W E G V F T
    StuI

2881 TAGGTACGA GGTACATGA ACACCCCGGG GCTTCCCGTG TGCCAGGACC ATCTTGAATT TTGGGAGGGC GTCTTTACAG
    ATCCGATGCT CGCATGTACT TGTGGGGCCC CGAAGGGCAC ACGTCTCTGG TAGAACTTAA AACCTCCCG CAGAAATGTC

+2  G L T H I D A H F L S Q T K Q S G E N L P Y L V A Y Q
    StuI
    ----

2961 GCCTCACTCA TATAGATGCC CACTTTCTAT CCCAGACAAA GCAGAGTGGG GAGAACCTTC CTTACCTGGT AGCTACCAA
    CGGAGTGAGT ATATCTACGG GTGAAAGATA GGTCTGTTT CGTCTCACC CTCTTGAAG GAATGGACCA TCGCATGGTT

+2  A T V C A R A Q A P P P S W D Q M W K C L I R L K P T
3041 GCCACCGTGT GCGCTAGGGC TCAAGCCCCT CCCCCATCGT GGGACACAGT GTGGAAGTGT TTGATTGCGC TCAAGCCCAC
    CGGTGGCACA CGCGATCCCG AGTTGGGGGA GGGGGTAGCA CCCTGGTCTA CACCTTCACA AACTAAGCGG AGTTGGGGTG
```

FIG. 3-5



pCMV-NS35

+2 L H G P T P L L Y R L G A V Q N E I T L T H P V T K  
3121 CCTCCATGGG CCAACACCCC TGCTATACAG ACTGGGGCT GTTCAGAATG AAATCACCCCT GACGCACCCA GTACACCAAT  
GGAGGTACCC GGTGTGGG ACGATATGTC TGACCCGCGA CAAGTCTTAC TTATGTGGA CTGCGTGGT CAGTGGTTTA

+2 Y I M T C M S A D L E V V T S T W V L V G G V L A A L  
3201 ACATCATGAC ATGCATGTC GCGACCTGG AGTCTGTAC GAGCACCTGG GTGCTCGTTG GCGGCGTCCT GGCTGCTTTG  
TGTAAGTCTG TAGGTACAGC CGGCTGGACC TCCAGCAGTG CTGCTGGACC CAGGAGCAAC GCGCGCAGGA CCGACGAAAC

+2 A A Y C L S T G C V V I V G R V V L S G K P A I I P D  
3281 GCGCGTATT GCCTGTCAAC AGGCTGGTG GTCATAGTGG GCAGGTCGT CTTGTCCGGG AAGCCGGCAA TCATACCTGA  
CGCGCATAA CGGACAGTG TCCGACGCAC CAGTATCACC CGTCCCAGCA GAACAGGCCC TTGCGCCGTT AGTATGGACT

+2 R E V L Y R E F D E M E E C S Q H L P Y I E Q G M M  
3361 CAGGGAAGTC CTCTACCGAG AGTTCGATGA GATGGAAGAG TGCTCTCAGC ACTTACCGTA CATCGAGCAA GGGATGATGC  
GTCCCTTCAG GAGATGGCTC TCAAGCTACT CTACCTTCTC ACGAGAGTGG TGAATGGCAT GTAGCTCGTT CCTACTACG

+2 L A E Q F K Q K A L G L L Q T A S R Q A E V I A P A V  
3441 TCGCGGAGCA GTTCAAGCAG AAGGCCCTCG GCCTCCTGCA GACCGGTCC CGTCAGGCAG AGTTATCGC CCCTGCTGTC  
AGCGGCTCGT CAAGTTCGTC TTCCGGGAGC CGGAGGACGT CTGGCGCAGG GCAGTCCGTC TCCAATAGCG GGGACGACAG

+2 Q T N W Q K L E T F W A K H M W N F I S G I Q Y L A G  
3521 CAGACCAACT GGCACCAACT CGAGACCTTC TGGGGAAGC ATATGTGGA CTTCATCAGT GGGATACAAT ACTTGGCGGG  
GTCTGGTTGA CCGTTTTTGA GCTCTGGAAG ACCGGTTG TATACACCTT GAAGTACTCA CCCTATGTTA TGAACCGCCC

+2 L S T L P G N P A I A S L M A F T A A V T S P L T T  
3601 CTGTCAACG CTGCCTGGTA ACCCGCCAT TGCTTCATTG ATGGCTTTTA CAGCTGTGT CACCAGCCCA CTAACCACTA  
GAACAGTTGC GACGGACCAT TGGGGCGGTA ACGAAGTAAC TACCGAAAAT GTCCAGGACA GTGTCGGGT GATTGGTGAT

FIG. 3-6





pCMV-NS35

+2 S Q T L L F N I L G G W V A A Q L A A P G A A T A F V  
3681 GCCAAACCCT CCTCTTCAAC ATATTGGGG GGTGGGTGC TGGCCAGCTC GCGGCCCGG GTGCCGCTAC TGCCTTTGTG  
CGGTTTGGGA GGAGAAGTIG TATAACCCCG CCACCCACCG ACGGTTCGAG CGCGGGGGC CACGGCGATG ACGGAAACAC

+2 G A G L A G A A I G S V G L G K V L I D I L A G Y G A  
3761 GCGCTGGCT TAGCTGGCG CGCCATCGG AGTGTGGAC TGGGAAGGT CCTCATAGAC ATCCTTGCAG GGTATGGCG  
CCGCGACCGA ATCGACCGG GCGGTAGCCG TCACAACCTG ACCCTTCCA GGAGTATCTG TAGGAACGTC CCATACCGCG

+2 G V A G A L V A F K I M S G E V P S T E D L V N L L  
3841 GCGGTGGCG GGAGCTCTTG TGGCATTCAA GATCATGAGC GGTGAGGTCC CCTCCACGGA GAACCTGGTC AATCTACTGC  
CCCGCACCGC CCTCGAGAAC ACCGTAAGTT CTAGTACTCG CCACTCCAGG GGAGTGCCT CCTGGACCAG TTAGATGACG

+2 P A I L S P G A L V V G V V C A A I L R R H V G P G E  
3921 CCGCCATCCT CTCGCCCGGA GCCCTCGTAG TCGCGTGGT CTGTGCAGCA ATACTGCGCC GGCACGTTGG CCCGGGGGAG  
GGCGTAGGA GAGCGGGCCT CCGGAGCATC AGCCGACCA GACAGTCTGT TATGACGCGG CCGTGAACC GGGCCCGCTC

+2 G A V Q W M N R L I A F A S R G N H V S P T H Y V P E  
4001 GGGCAGTGC AGTGGATGAA CCGGCTGATA GCCTTGCCT CCCGGGGAA CCATGTTTCC CCCACGCACT ACGTGCCGGA  
CCCCGTACG TCACCTACTT GGCGACTAT CGGAAGCGGA GGGCCCCCTT GGTACAAAGG GGTGCGTGA TGCACGGCCT

+2 S D A A A R V T A I L S S L T V T Q L L R R L H Q W  
4081 GAGCGATGCA GCTGCCCGG TCACTGCCAT ACTCAGCAGC CTCACTGTAA CCCAGCTCCT GAGGCGACTG CACCACTGGA  
CTCGTACGT CGACGGGGCG AGTGACGTA TGATCGTCG GAGTGACATT GGTTCGAGGA CTCCGCTGAC GTGGTCACT

+2 I S S E C T T P C S G S W L R D I W D W I C E V L S D  
4161 TAAGCTCGGA GTGTACCACT CCATGCTCCG GTTCTGGCT AAGGACATC TGGGACTGGA TATGCGAGGT GTTGAGCGAC  
ATTGAGCCT CACATGGTGA GGTACGAGG CAAGGACCGA TTCCCTGTAG ACCCTGACCT ATACGCTCCA CAACTCGCTG

FIG. 3-7



pCMV-NS35

+2 F K T W L K A K L M P Q L P G I P F V S C Q R G Y K G  
BamHI  
-----  
4241 TTAAAGACCT GGCTAAAAGC TAAGCTCATG CCACAGCTGC CTGGGATCCC CTTTGTGTCC TGCCAGCGCG GGTATAAGGG  
AAATTCTGGA CCGATTCTCG ATTGAGTAC GGTGTCGACG GACCCTAGGG GAAACACAGG ACGTGGCGC CCATATTCCC  
-----  
+2 V W R G D G I M H T R C H C G A E I T G H V K N G T  
4321 GGTCTGGGA GGGACGGCA TCATGCACAC TCGTGGCAGC TGTGGAGCTG AGATCACTGG ACATGTCAA AACGGGACGA  
CCAGACCGCT CCCCTGCCGT AGTACGTGTG AGCGACGGTG ACACCTCGAC TCTAGTGACC TGTACAGTTT TTGCCCTGCT  
-----  
+2 M R I V G P R T C R N M W S G T F P I N A Y T T G P C  
4401 TGAGGATCGT CGTCTTAGG ACCTGCAGGA ACATGTGGAG TGGACCTTC CCCATTAAATG CCTACACCAC GGGCCCTGT  
ACTCTAGCA GCCAGGATCC TGGACGTCTT TGTACACCTC ACCCTGGAAG GGGTAATTAC GGATGTGGTG CCGGGGACAC  
-----  
+2 T P L P A P N Y T F A L W R V S A E E Y V E I R Q V G  
4481 ACCCCCTTC CTGGCCGAA CTACACCTTC GCGTATGGA GGTGTCTGC AGAGGAATAC GTGGAGATAA GGCAGGTGG  
TGGGGGAAG GACGGGGCTT GATGTGCAAG CCGGATACCT CCCACAGAGC TCTCTTATG CACCTCTATT CCGTCCACCC  
-----  
+2 D F H Y V T G M T T D N L K C P C Q V P S P E F F T  
4561 GGACTTCCAC TACGTGACGG GTATGACTAC TGACAATCTT AAATGCCGT GCCAGGTCCC ATGCCCGAA TTTTTCACAG  
CCTGAAGCTG ATGCACTGCC CATACTGATG ACTGTAGAA TTTACGGGA CCGTCCAGGG TAGCGGGCTT AAAAAGTCTC  
-----  
+2 E L D G V R L H R F A P P C K P L L R E E V S F R V G  
4641 AATTGGACGG GGTGGGCTA CATAGTTTG CGCCCCCTG CAAGCCCTTG CTGCGGGAGG AGGTATCATT CAGAGTAGGA  
TTAACCTGCC CCACGGGAT GTATCCAAAC GCGGGGGAC GTTCGGGAAC GACGCCCTCC TCCATAGTAA GTCTCATCTT  
-----  
+2 L H E Y P V G S Q L P C E P E P D V A V L T S M L T D  
4721 CTCCACGAAT ACCCGGTAGG GTCCGAATTA CCTTGGGAGC CCGAACCGGA CGTGGCCGTG TTGACGTCCA TGCTCACTGA  
GAGGTGCTTA TGGGCCATCC CAGCGTTAAT GGAACGGCTG GCTTGGCTT GCACCGGCAC AACTGCAGGT ACGAGTGACT  
-----

FIG. 3-8



pCMV-NS35

+2 P S H I T A E A A G R R L A R G S P P S V A S S A  
4801 TCCCTCCCAT ATAACAGCAG AGCGGGCGG GCGAAGTTG GCGAGGGGAT CACCCCCCTC TGTCGGCAGC TCCTCGGCTA  
AGGAGGGTA TATTGTCGTC TCCGCGGGC CGCTTCCAAC CGTCCCCTA GTGGGGGAG ACACGGTGC AGGAGCCGAT

---

+2 S Q L S A P S L K A T C T A N H D S P D A E L I E A N  
4881 GCCAGCTATC CGCTCCATCT CTCAAGGCAA CTTCACCGC TAACCATGAC TCCCCTGATG CTGAGCTCAT AGAGGCCAAC  
CGGTCGATAG CCGAGGTAGA GAGTTCCGTT GAACGTGGC ATTGTACTG AGGGACTAC GACTCGAGTA TCCTCGGTTG

---

+2 L L W R Q E M G G N I T R V E S E N K V V I L D S F D  
4961 CTCCTATGGA GGCAGGAGAT GGGCGGCAAC ATCACCAGG TTGAGTCAGA AAACAAAGT GTGATTCTGG ACTCCTTCGA  
GAGGATACCT CCGTCTCTA CCCGCCGTTG TAGTGTCCC AACTCAGTCT TTTGTTTTCAC CACTAAGACC TGAGGAAGCT

---

+2 P L V A E E D E R E I S V P A E I L R K S R R F A Q  
5041 TCCGCTGTG GCGAGGAGG ACGAGCGGA GATCTCCGTA CCGCAGAAA TCCTCGGAA GTCTCGGAGA TTCGCCCAGG  
AGGCGAACAC CGCTCCTCC TGCTCGCCCT CTAGAGGAT GGGCGTCTT AGGACGCCT CAGAGCCTCT AAGCGGTCC

---

+2 A L P V W A R P D Y N P P L V E T W K K P D Y E P P V  
5121 CCCTGCCCCGTTGGCGCGG CCGGACTATA ACCCCCGCT AGTGAGACG TGGAAGAGC CCGACTACGA ACCACCTGTG  
GGGAGGGCA AACCGCGCGG GGCCTGATAT TGGGGGGCGA TCACCTCTGC ACCTTTTTCG GGCTGATGCT TGGTGGACAC

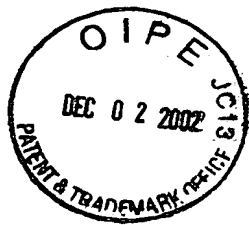
---

+2 V H G C P L P P P K S P P V P P P R K K R T V V L T E  
5201 GTCCATGGCT GCCCGCTCC ACCTCCAAAG TCCCCTCCTG TGCCTCCGC TCAGGAGAGC CGGACGGTGG TCCTCACTGA  
CAGGTACCGA CCGGCGAAGG TGGAGTTTC AGGGAGGAC ACGGAGCGG AGCTTCTTC GCCTGCCACC AGGAGTGA

---

+2 S T L S T A L A E L A T R S F G S S S T S G I T G D  
5281 ATCAACCCTA TCTACTGCCT TGGCCGAGCT CGCCACCAGA AGCTTTGGCA GCTCCTCAAC TTCGGGCAAT ACGGGCGACA  
TAGTTGGAT AGATGACGGA ACCGGCTCGA CCGGTGGTCT TCGAAACCGT CGAGAGTTG AAGGCCGTAA TGCCCGCTGT

FIG. 3-9



pCMV-NS35

+2 N T T T S S E P A P S G C P P D S D A E S Y S S M P P  
5361 ATACGACAAC ATCCTCTGAG CCGGCCCTTT CTGGCTGCC CCGGACTCC GACGCTGAGT CCTATTCTC CATGCCCCC  
TATGCTGTTG TAGGAGACTC GGGCGGGGAA GACCGACGG GGGGCTGAGG CTGGGACTCA GGATAAGGAG GTACGGGGGG

+2 L E G E P G D P D L S D G S W S T V S S E A N A E D V  
BamHI

5441 CTGGAGGGG AGCCTGGGA TCCGGATCTT AGCGACGGT CATGGTCAAC GGTCACTAGT GAGGCCAACG CGGAGGATGT  
GACCTCCCC TCGGACCCCT AGGCCTAGAA TCGCTGCCC GTACCACTTG CCAGTCATCA CTCGGGTTC GCCTCCTACA

+2 V C C S M S Y S W T G A L V T P C A A E E Q K L P I  
5521 CGTGTGCTGC TCAATGTCTT ACTCTTGGAC AGGGCACTC GTCACCCCGT GCGCGCGGA AGAACAGAAA CTGCCCCATCA  
GCACACGACG AGTTACAGAA TGAGAACCTG TCCGCTGAG CAGTGGGCA GCGGGGCGCT TCTTGCTTT GACGGGTAGT

+2 N A L S N S L L R H H N L V Y S T T S R S A C Q R Q K  
5601 ATGCACTAAG CAACCTGTTG CTACGTACC ACAATTGGT GTATTCCACC ACCTCAGCA GTGCTTGCCA AAGGCAGAAG  
TACGTGATTC GTTGAAGCAAC GATGCAGTGG TGTTAAACCA CATAAGGTGG TGGAGTGCCT CACGAACGGT TTCCGTCTTC

+2 K V T F D R L Q V L D S H Y Q D V L K E V K A A S K  
5681 AAAGTCACAT TTGACAGACT GCAAGTTCTG GACAGCCATT ACCAGGACGT ACTCAAGGAG GTTAAAGCAG CGGCGTCAAA  
TTTCAGTGTA AACTGTCTGA CGTTCAAGAC CTGTGCTGTA TGGTCTGCA TGAGTTCTC CAATTTCTC GCGGCAGTTT

+2 V K A N L L S V E E A C S L T P P H S A K S K F G Y  
5761 ACTGAAGGCT AACTTGCTAT CCGTAGAGGA AGCTTGCAGC CTGACGCCCC CACACTCAGC CAAATCCAAG TTGTTATG  
TCACTTCCGA TTGAACGATA GGCATCTCCT TCGAACGTGG GACTGGGGG GTGTGAGTGG GTTAGGTTT AAACCAATAC

FIG. 3-10



PCMV-NS35

+2 G A K D V R C H A R K A V T H I N S V W K D L L E D N  
5841 GGGCAAAGA CGTCGGTTC CATGCCAGAA AGCCGTAAC CCACATCAAC TCCGTGTGA AAGACCTTCT GGAAGACAAT  
CCCGTTTCT GCAGGCAACG GTACGGTCTT TCCGGCATG GGTAGTTG AGGCACACCT TTCTGGAAGA CCTTCTGTTA

+2 V T P I D T T I M A K N E V F C V Q P E K G G R K P A  
5921 GTAACACCAA TAGACACTAC CATCATGGCT AAGAACGAGG TTTTCTGCGT TCAGCCTGAG AAGGGGGTC GTAAGCCAGC  
CATTCTGGTT ATCTGTGATG GTAGTACCGA TTCTTGCTCC AAAAGACGCA AGTCGGACTC TTCCCCCAG CATTGGGTGG

+2 R L I V F P D L G V R V C E K M A L Y D V V T K L P  
6001 TCGTCTCATC GTGTTCCCG ATCTGGCGT GCGCGTGTG GAAAGATGG CTTTGTACGA CGTGGTTACA AAGCTCCCCT  
AGCAGAGTAG CACAAGGGGC TAGACCCGCA CGCGCACACG CTTTCTTACC GAAACATGCT GCACCAATGT TTCGAGGGGA

+2 L A V M G S S Y G F Q Y S P G Q R V E F L V Q A W K S  
EcoRI

6081 TGGCCGTGAT GGAAGCTCC TACGGATTCC AATACTACC AGGACAGCGG GTTGAATTCC TCGTGCAAGC GTGGAAGTCC  
ACCGGCACTA CCCTTCGAGG ATGCTAAGG TTATGAGTGG TCCTGTGCGC CAACCTAAGG AGCACGTTCC CACCTTCAGG

+2 K K T P M G F S Y D T R C F D S T V T E S D I R T E E  
6161 AAGAAACCC CAATGGGGTT CTCGTATGAT ACCCGTGTCT TTGACTCCAC AGTCACCTGAG AGGCACATCC GTACGGAGGA  
TTCTTTTGGG GTTACCCCAA GACATACTA TGGCGACGA AACTGAGGTG TCAGTGACTC TCGCTGTAGG CATGCCTCCT

+2 A I Y Q C C D L D P Q A R V A I K S L T E R L Y V G  
6241 GGCAATCTAC CAATGTTGTG ACCTCGACCC CCAAGCCGC GTGGCCATCA AGTCCCTCAC CGAGAGGCTT TATGTTGGGG  
CCGTAGATG GTTACAACAC TGGAGCTGGG GGTTCGGGG CACCGGTAGT TCAGGGAGTG GCTCTCCGAA ATACAACCCC

+2 G P L T N S R G E N C G Y R R C R A S G V L T T S C G  
6321 GCCCTCTTAC CAATTCAAGG GGGGAGAACT GCGGCTATCG CAGGTGCGC GCGAGGGCG TACTGACAAC TAGCTGTGGT  
GGGGAAGATG GTTAAGTTCC CCCCTCTTGA CGCGATAGC GTCCACGGCG CGCTCGCCG ATGACTGTG ATCGACACCA

FIG. 3-11



pCMV-NS35

+2 N T L T C Y I K A R A A C R A A G L Q D C T M L V C G  
6401 AACACCCCTCA CTTGCTACAT CAAGGCCCGG GCAGCCTGTC GAGCCG CAGG GCTCCAGGAC TGCACCATGC TCGTGTCTGG  
TTGTGGGAGT GAACGATGTA GTTCCGGGCC CGTCGGACAG CTCGGCGTCC CGAGGTCTG ACGTGGTAGG AGCACACACC

---

+2 D D L V V I C E S A G V Q E D A A S L R A F T E A M  
6481 CGACGACTTA GTCGTTATCT GTGAAAGCGC GGGGTCCAG GAGGACGCGG CGAGCCTGAG AGCCTTCACG GAGGCTATGA  
GCTGCTGAAT CAGCAATAGA CACTTTCGCG CCCCCAGGTC CTCCTGGCGC GCTCGGACTC TCGGAAGTGC CTCCGATACT

---

+2 T R Y S A P P G D P P Q P E Y D L E L I T S C S S N V  
6561 CCAGGTACTC CGCCCCCCT GGGACCCCG CACAACCGA ATACGACTTG GAGCTCATAA CATCATGCTC CTCCAACGTG  
GGTCCATGAG GCGGGGGGA CCCCTGGGG GTGTGGTCT TATGCTGAAC CTCGAGTATT GTAGTACGAG GAGGTTCAC

---

+2 S V A H D G A G K R V Y Y L T R D P T T P L A R A A W  
6641 TCAGTCGCC ACGACGGCGG TGAAAGAGG GTCTACTACC TCACCCGTGA CCCTACAACC CCCCTCGGA GAGCTGCGTG  
AGTCAGCGG TGCTGCCGG ACCTTTCTCC CAGATGATGG AGTGGGCACT GGGATGTTGG GGGGAGCGCT CTCGACGAC

---

+2 E T A R H T P V N S W L G N I I M F A P T L W A R M  
6721 GGAGACAGCA AGACACACTC CAGTCAATTC CTGGCTAGGC AACATAATCA TGTTCGCCCC CACACTGTGG GCGAGGATGA  
CCTCTGCGT TCTGTGTGAG GTCAGTTAAG GACCGATCCG TTGTATTAGT ACAACGGGG GTGTGACACC CGCTCCTACT

---

+2 I L M T H F F S V L I A R D Q L E Q A L D C E I Y G A  
6801 TACTGATGAC CCATTTCCTT AGCGTCCTTA TAGCCAGGA CCAGCTTGAA CAGGCCCTCG ATTGGGAGAT CTACGGGGCC  
ATGACTACTG GGTAAAGAAA TCGCAGGAAT ATCGTCCCT GGTGCAACTT GTCCGGGAGC TAACGCTCTA GATGCCCCGG

---

+2 C Y S I E P L D L P P I I Q R L H G L S A F S L H S Y  
6881 TGCTACTCCA TAGAACCACT GGATCTACCT CCAATCAATC AAAGACTCCA TGGCCTCAGC GCATTTTCAC TCCACAGTTA  
ACGATGAGGT ATCTTGGTGA CCTAGATGA GGTAGTAGG TTTCTGAGT ACCGGAGTCC CGTAAAGTG AGGTGTCAAT

FIG. 3-12



pCMV-NS35

+2 S P G E I N R V A A C L R K L G V P P L R A W R H R  
6961 CTCTCCAGGT GAAATCAATA GGGTGGCCGC ATGCCTCAGA AAACCTGGGG TACCGCCCTT GCGAGCTTGG AGACACCGGG  
GAGAGTCCA CTTTAGTTAT CCCACCGGGG TACGGAGTCT TTTGAACCCC ATGGCGGAA CGCTCGAACC TCTGTGGCCC

+2 A R S V R A R L L A R G G R A A I C G K Y L F N W A V  
7041 CCCGGAGCGT CCGCGCTAGG CTTCTGGCCA GAGGAGGCAG GGCTGCCATA TGTGGCAAGT ACCTCTTCAA CTGGGCAGTA  
GGGCTCGCA GCGCGGATCC GAAGACCGGT CTCTCCGTC CCGACGGTAT ACACCGTTCA TGGAGAAATT GACCCGTCAT

+2 R T K L K L T P I A A A G Q L D L S G W F T A G Y S G  
7121 AGAACAAAGC TCAAACCTCAC TCCAATAGCG GCGCTGGCC AGCTGGACTT GTCCGGCTGG TTCACGGCTG GCTACAGCGG  
TCTTGTTTCG AGTTGAGTG AGTTATCGC CCGCGACCGG TCGACCTGAA CAGGCCGACC AAGTCCGAC CGATGTGGCC

+2 G D I Y H S V S H A R P R W I W F C L L L L A A G V  
7201 GGGAGACATT TATCACAGCG TGTCTCATGC CCGGCCCGCG TGGATCTGGT TTTGCCTACT CCTGCTTGCT GCAGGGGTAG  
CCCTCTGTAA ATAGTGTCC ACAGAGTACG GGCCCGGGCG ACCTAGACCA AACGGATGA GGACGAACGA CGTCCCCATC

+2 G I Y L L P N R  
7281 GCATCTACCT CCTCCCCAAC CGATGAAGGT TGGGGTAAAC ACTCCGGCCT AAAAAAAAAA AAAAAATCTAG AAAGCGCGCG  
CGTAGATGGA GGAGGGGTTG GCTACTTCCA ACCCATTTG TGAGGCCGGA TTTTTTTTTT TTTTITAGATC TTTCGGCGCG

BAMHI MluI  
-----

7361 CAAGATATCA AGGATCCACT ACGCGTTAGA GCTCGCTGAT CAGCCTCGAC TGTGCCTTCT AGTTGCCAGC CATCTGTTGT  
GTTCTATAGT TCCTAGGTGA TCGGCAATCT CGAGCGACTA GTCGGAGCTG ACACGGAAGA TCAACGGTGC GTAGACAACA

7441 TTGCCCCCTCC CCGGTGCCTT CCTTGACCCCT GGAAGGTGCC ACTCCCCACTG TCCTTTCCTA ATAAATGAG GAAATTGCAT  
AACGGGGAGG GGGCACGGAA GGAAC TGGA GGAAGGTGAC AGGAAAGGAT TATTTTACTC CTTTAAACGTA

FIG. 3-13



pCMV-NS35

7521 CGCATTGTCT GAGTAGGTGT CATTCTATTTC TGGGGGTGG GGTGGGGCAG GACAGCAAGG GGGAGGATTG GGAAGACAAT  
GCGTAACAGA CTCATCCACA GTAAGATAAG ACCCCCCACC CCACCCCGTC CTGTGCTTCC CCCTCCTAAC CCTTCTGTTA

---

7601 AGCAGGCATG CTGGGGAGCT CTTCGGCTTC CTGCGTCACT GACTCGCTGC GCTCGGTCTG TCGGCTGCGG CGAGCGGTAT  
TCGTCCGTAC GACCCCTCGA GAAGGCGAAG GAGCGAGTGA CTGAGCGACG CGAGCCAGCA AGCGGACGCC GCTCGCCATA

---

7681 CAGCTCACTC AAAGGCGGTA ATACGGTTAT CCACAGAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG  
GTCGAGTGAG TTTCCGCCAT TATGCCAATA GGTGCTTAG TCCCCATTG CGTCCCTTCT TGTACACTCG TTTTCCGGTC

---

7761 CAAAAGGCCA GGAACCGTAA AAAGGCCGCG TTGCTGGCT TTTTCCATAG GCTCCGCCCC CCTGACGAGC ATCACAATAA  
GTTTCCGGT CCTTGGCATT TTTCCGGGCG AAGCACCGCA AAAAGGTATC CGAGGCGGG GGACTGCTCG TAGTGTTTTT

---

7841 TCGACGCTCA AGTCAGAGGT GCGGAAACCC GACAGGACTA TAAAGATACC AGGCGTTTCC CCCTGGAAGC TCCCTCGTGC  
AGCTGCGACT TCAGTCTCCA CCGCTTTGGG CTGTCTGAT ATTTCTATGG TCCGCCAAAG GGGACCTTCC AGGAGCACG

---

7921 GCTCTCCTGT TCCGACCCTG CCGCTTACCG GATACCTGTC CGCTTCTC CTTCGGGAA GCGTGGCGCT TTCTCAATGC  
CGAGAGGACA AGCTGGGAC GCGCAATGGC CTATGGACAG GCGGAAAGAG GGAAGCCCTT CGCACCGCGA AAGAGTTACG

---

8001 TCACGCTGTA GGTATCTCAG TTCGGTGTAG GTCGTTCCGT CCAAGCTGG CTGTGTGCAC GAACCCCGCG TTCAGCCCGA  
AGTGGACAT CCATAGAGTC AAGCACATC CAGCAAGCGA GGTTCGACCC GACACACGTG CTTGGGGGGC AAGTCGGGCT

---

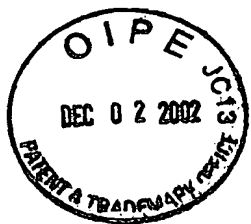
8081 CCGCTCGCC TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG  
GGCGACGGG AATAGGCCAT TGATAGCAGA ACTCAGGTTG GGCAATTCTG TGCTGAATAG CCGTGACCGT CGTCGGTGAC

---

8161 GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG AAGTGGTGGC CTAACACTAG CTACACTAGA  
CATTGTCTTA ATGCTCTCGC TCCATACATC CGCCACGATG TCTCAAGAAC TTCACCACCG GATTGATGCC GATGTGATCT

FIG. 3-14





pCMV-NS35

8241	AGGACAGTAT TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAAAACA TCCTGTGCATA AACCATAGAC GCGAGACGAC TTCGGTCAAT GGAAGCCCTT TTCTCAACCA TCGAGAACCTA GGCCGTTTGT
8321	AACCACCGCT GGTAGCGGTG GTTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT TTGGTGGCGA CCATCGCCAC CAAAAAACA AACGTTGCTC GTCTAATGCG CGTCTTTTTT TCCTAGAGTT CTCTAGGAA
8401	TGATCTTTTC TACGGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA GGGATTTTGG TCATGAGATT ATCAAAAAGG ACTAGAAAAG ATGCCCCAGA CTGCGAGTCA CCTTGCTTTT GAGTGCAATT CCTAAAAACC AGTACTCTAA TAGTTTTTCC
8481	ATCTTCACCT AGATCCCTTTT AAATTAAAAA TGAAGTTTTA AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG TAGAAGTGGA TCTAGGAAAA TTTAATTTTTT ACTTCAAAAT TTAGTTAGAT TTCATATATA CTCATTTGAA CCAGACTGTC
8561	TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGCTATTTC GTTCATCCAT AGTTGCCTGA CTCCCCGTGC AATGGTTACG AATTAGTCAC TCCGTGGATA GAGTCGCTAG ACAGATAAAG CAAGTAGGTA TCAACGGACT GAGGGGCAGC
8641	TGTAGATAAC TACGATACGG GAGGGCTTAC CATCTGGCCC CAGTGTGCA ATGATACCGC GAGACCCACG CTCACCGGCT ACATCTATTG ATGCTATGCC CTCCCGAATG GTAGACCGGG GTCACGACGT TACTATGGCG CTCTGGGTGC GAGTGGCCGA
8721	CCAGATTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG AGGCAGAAAG TGGTCCTGCA ACTTTATCCG CCTCCATCCA GGTCTAAATA GTCGTTATTT GGTCCGTCCG CCTTCCCGGC TCGCGTCTTC ACCAGGACGT TGAATAGGC GGAGGTAGGT
8801	GTCTATTAT TGTTCGCGG AAGCTAGAGT AAGTAGTTCG CCAGTTAATA GTTTGCGCAA CGTTGTGTC ATTGCTACAG CAGATAATTA ACAACGGCCC TTCGATCTCA TTCAATCAAGC GGTCATTTAT CAAACGCGTT GCAACAAACGG TAACGATGTC
8881	GCATCGTGGT GTCACGCTCG TCGTTTGTA TGGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGCGGAGT TACATGATCC CGTAGCACCA CAGTGGGAGC AGCAAACCAT ACCGAAGTAA GTCGAGGCCA AGGTTGCTA GTTCCGCTCA ATGTACTAGG

FIG. 3-15



pCMV-NS35

8961 CCCATGTTGT GCAAAAAAGC GGTAGCTCC TTCGGTCCTC CGATCGTTGT CAGAAGTAAG TTGGCCGCAG TGTTATCACT  
GGGTACAACA CGTTTTTTTCG CCAATCGAGG AAGCAGGAG GCTAGCAACA GTCTTCATTC AACCGGCGTC ACAATAGTGA

---

9041 CATGGTTATG GCAGCACTGC ATAAATCTCT TACTGTGATG CCATCCGTAA GATGCTTTTC TGTGACTGGT GAGTACTCAA  
GTACCAATAC CGTCGTGACG TATTAAGAGA ATGACAGTAC GGTAGGCATT CTACGAAAAG ACAC TGACCA CT CATGAGTT

---

9121 CCAAGTCATT CTGAGAAATAG TGTATGCGGC GACCGAGTTG CTCTTGCCCG GCGTCAATAC GGGATAATAC CGCGCCACAT  
GGTTCAGTAA GACTCTTATC ACATACGCCG CTGGCTCAAC GAGAACGGGC CGCAGTTATG CCTATTATG GCGCGGTGTA

---

9201 AGCAGAACTT TAAAAGTGCT CATCATTTGA AAACGTTCTT CGGGGCGGAA ACTCTCAAGG ATCTTACCGC TGTTGAGATC  
TCGTCTTGAA ATTTTCACGA GTAGTAACCT TTTCGAAGAA GCCCGGCTTT TGAGAGTTCC TAGAATGGCG ACAACTCTAG

---

9281 CAGTTCGATG TAACCCACTC GTGCACCCAA CTGATCTTCA GCATCTTTTA CTTTCACCCAG CGTTTCTGGG TGAGCAAAAA  
GTCAAGCTAC ATTGGGTGAG CACGTGGGTT GACTAGAAGT CGTAGAAAAAT GAAAGTGGTC GCAAAGACCC ACTCGTTTTT

---

9361 CAGGAAGGCA AAATGCCGCA AAAAAAGGAA TAAGGGCGAC ACGAAATGT TGAATACTCA TACTCTTCCT TTTTCAATAT  
GTCCCTCCGT TTTACGGCGT TTTTCCCTT ATTCCCGCTG TGCCTTACA ACTTATGAGT ATGAGAAGGA AAAAGTTATA

---

9441 TATTGAAGCA TTTATCAGG TTATTGTCTC ATGAGCGGAT ACATATTTGA ATGTATTTAG AAAAAATAAC AAATAGGGGT  
ATAACTTCGT AAATAGTCCC AATAACAGAG TACTCGCCTA TGTATAAACT TACATAAATC TTTTATTG TTTATCCCCA

---

9521 TCCGGGCACA TTTCCCCGAA AAGTGCCACC TGACGTCTAA GAAACCATTA TTATCATGAC ATTAACCTAT AAAAATAGGC  
AGGCGGTGT AAAGGGCTT TTCACGGTGG ACTGCAGATT CTTTGGTAAT AATAGTACTG TAATTGGATA TTTTATCCG

---

9601 GTATCAGCAG GCCCTTTCGT C  
CATAGTGCTC CGGAAAGCA G

FIG. 3-16

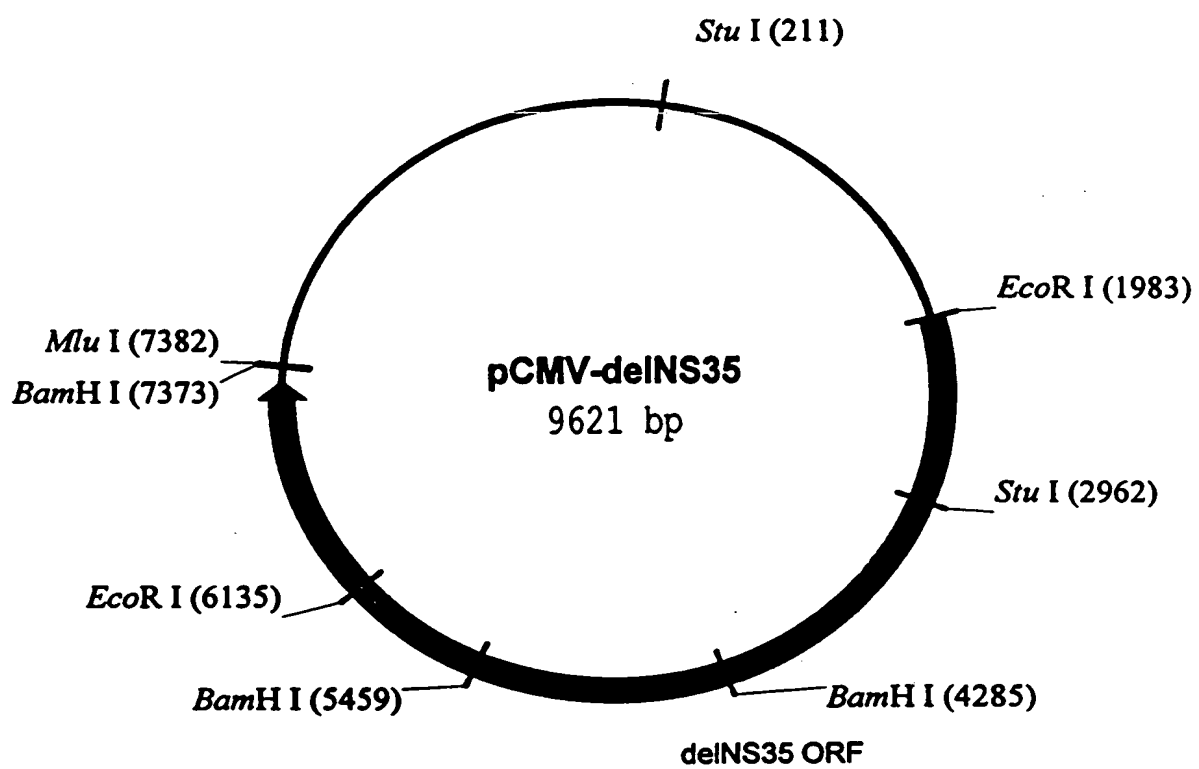


FIG. 4



pCMV-deINS35

1	TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGCT GTAAAGCGGAT AGCGCGCAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGCG CTCTGCCAGT CTCGAACAGA CATTGGCCTA
81	GCCGGGAGCA GACAAGCCCG TCAGGGGCGG TCAGGGGGTG TTGGCGGGTG TCGGGGGTGG CTTAACATATG CGGCATCAGA CGGCCCTCGT CTGTTGGGC AGTCCCGGC AGTCCCGGC AACCGCCAC AGCCCGACC GAATTGATAC GCCGTAGTCT
161	GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTGGCA AAAGCCTAGG CCTCCAAAA AGCCTCCTCA CTACTTCTGG CGTCTAACAT GACTCTCAG TGGTACTT CGAAAAACGT TTTCGGATCC GGAGTTTTT TCGGAGGAGT GATGAAGACC
241	AATAGCTCAG AGCCGAGGC GGCCTCGGC TCTGCATAAA TAAAAAAAT TAGTCAGCCA TGGGGCGGAG ATTGGGCGGA TTATCGAGTC TCCGGCTCCG CCGGAGCCG AGAGTATTT ATTTTTTTA ATCAGTCGGT ACCCGCCTC TTACCCGCCT
321	ACTGGGCGG GAGGGAATTA TTGGCTATTG GCCATTGCAT ACCTTGATC TATATCATAA TATGTACATT TATATTGGCT TGACCCGCC CTCCCTTAAT AACCGATAAC CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA
401	CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGTC ATTAGTTTCA GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAAGTA
481	AGCCCATATA TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCGCCCAT TCGGGTATAT ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GCGCGGTAA
561	GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTACGGT CTGCAGTTAT TACTGCATAC AAGGTATCA TTGGGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA

StuI

-----

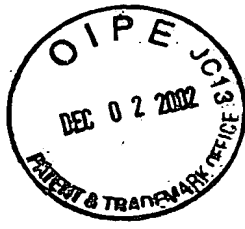
FIG. 5-1



pCMV-delNS35

641	AAACTGCCCA	CTTGGCAGTA	CATCAAGTGT	ATCATATGCC	AAGTCGGCCC	CCTATTGACG	TCAATGACGG	TAAATGGCCC
	TTTGACGGGT	GAACCGTCAT	GTAGTTCACA	TAGTATACGG	TTCAGGGGGG	GGATAACTGC	AGTTACTGCC	ATTTACCGGG
721	GCCTGGCATT	ATGCCCAGTA	CATGACCTTA	CGGGACTTTC	CTACTTGGCA	GTACATCTAC	GTATTAGTCA	TCGCTATTAC
	CGGACCGTAA	TACGGGTCAT	GTACTGGAAT	GCCCTGAAAG	GATGAACCGT	CATGTAGATG	CATAATCAGT	AGCGATAATG
801	CATGGTGATG	CGGTTTTTGGC	AGTACACCAA	TGGGCGTGGA	TAGCGGTTTG	ACTCACGGGG	ATTTCCAAGT	CTCCACCCCA
	GTACCACTAC	GCCAAAACCG	TCAATGTGTT	ACCCGCACCT	ATCGCCAAAC	TGAGTGCCCC	TAAAGGTTCA	GAGGTGGGT
881	TTGACGTCAA	TGGGAGTTTG	TTTTTGGCACC	AAAATCAACG	GGACTTTTCCA	AAATGTCTGA	ATAACCCCGC	CCC GTTGACG
	AACTGCAGTT	ACCCTCAAAC	AAAACCGTGG	TTTTTAGTTGC	CCTGAAAGGT	TTTACAGCAT	TATTGGGGCG	GGGCAACTGC
961	CAAAATGGGG	GTAGGCGTGT	ACGGTGGGAG	GTCTATATAA	GCAGAGCTCG	TTTAGTGAAC	CGTCAGATCG	CCTGGAGACG
	CTTTACCCGC	CATCCGCACA	TGCCACCCTC	CAGATATATT	CGTCTCGAGC	AAATCACTTG	GCAGTCTAGC	GGACCTCTGC
1041	CCATCCACGC	TGTTTTTGACC	TCCATAGAAG	ACACCGGGAC	CGATCCAGCC	TCCGCGGGCG	GGAACGGTGC	ATTGGAACGC
	GGTAGGTGCG	ACAAAACCTGG	AGGTATCTTC	TGTGGCCCTG	GCTAGGTGCG	AGCGCCCGGC	CCTTGCCACG	TAACTTTGCG
1121	GGATTCCCGG	TGCCAAGAGT	GACGTAAGTA	CGGCCTATAG	ACTCTATAGG	CACACCCCTT	TGGCTCTTAT	GCATGCTATA
	CCTAAGGGGC	ACGGTTCTCA	CTGCATTTCAT	GGCGGATATC	TGAGATATCC	GTGTGGGGAA	ACCGAGAATA	CCTACCGATAT
1201	CTGTTTTTTGG	CTTGGGGCCT	ATACACCCGC	GCTCCTTATG	CTATAGGTGA	TGGTATAGCT	TAGCCTATAG	GTGTGGGTTA
	GACAAAACC	GAACCCCGGA	TATGTGGGG	CGAGGAATAC	GATATCCACT	ACCATATCGA	ATCGGATATC	CACACCCAAAT
1281	TTGACCATTAA	TTGACCACTC	CCCTATTGGT	GACGATACTT	TCCATTACTA	ATCCATAACA	TGGCTCTTTG	CCACAACCTAT
	AACTGGTAAT	AACTGGTGAG	GGGATAACCA	CTGCTATGAA	AGGTAATGAT	TAGGTATTGT	ACCGAGAAAC	GCTGTTGATA

FIG. 5-2



pCMV-delINS35

1361	CTCTATTGGC TATATGCCAA TACTCTGTCC TTACAGAGACT GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTAT	
	GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCCCTGA GACATAAAAA TGTCTTACCC CAGGTAAATA	
1441	TATTTACAAA TTCACATATA CAACAACGCC GTCCCCCGTG CCGGCAGTTT TTATTAAACA TAGCGTGGGA TCTCCGACAT	
	ATAAATGTTT AAGTGATAT GTTGTTCGG CAGGGGGCAC GGGCGTCAAA AATAATTGT ATCGCACCCCT AGAGGCTGTA	
1521	CTCGGGTACG TGTTCGGGAC ATGGGCTCTT CTCCGGTAGC GCGGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA	
	GAGCCCATGC ACAAGGCCCTG TACCCGAGAA GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT	
1601	GCGGCTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA CTTAGGCACA GCACAATGCC CACCACCACC	
	CGCCGAGTAC CAGCGAGCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT GAATCCGTGT CGTGTACCG GTGGTGGTGG	
1681	AGTGTGCGGC ACAAGGCCGT GCGGTTAGG TATGTGTCG AAATGAGCT CGGAGATTGG GCTCGCACCT GGACGCAGAT	
	TCACACGGCG TGTTCGGCA CCGCCATCCC ATACACAGC TTTTACTCGA GCCTCTAACC CGAGCGTGA CTTGCGTCTA	
1761	GGAAGACTTA AGCAGCGGC AGAAGAAGAT GCAGGCAAGT GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT	
	CCTTCTGAAT TCCGTGCGCG TCTTCTTCTA CGTCCGTCCA CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA	
1841	TGCGGTGCTG TTAACGGTGG AGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG CGCCACCAGA CATAATAGCT	
	ACGCCACGAC AATTGCCACC TCCGCTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC GCGGTGGTCT GTATTATCGA	
+2		M A A
		EcoRI
		----
1921	GACAGACTAA CAGACTGTTT CTTTCCATGG GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTACCC ATGGCTGCAT	
	CTGTCTGATT GTCTGACAAG GAAAGGTACC CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTGG TACCGACGTA	

FIG. 5-3



pCMV-delINS35

+2 Y A A Q G Y K V L V L N P S V A A T L G F G A Y M S K  
2001 ATGCAGCTCA GGGCTATAAG GTGCTAGTAC TCAACCCCTC TCTTGCTGCA ACACTGGGT TTGGTGCTTA CATGTCCAAG  
TAGCTCGAGT CCGGATATTC CAGGATCATG AGTTGGGAG ACAACGACGT TGTGACCCGA AACCAACGAAT GTACAGGTTT

+2 A H G I D P N I R T G V R T I T T G S P I T Y S T Y G  
2081 GCTCATGGGA TCGATCCTAA CATCAGGACC GGGGTGAGAA CAATTACCAC TGGCAGCCCC ATCAGCTACT CCACCTACGG  
CGAGTACCCT AGCTAGGATT GTAGTCCTGG CCCCACTCTT GTTAATGCTG ACCGTGGGG TAGTGCATGA GGTGGATGCC

+2 K F L A D G G C S G G A Y D I I I C D E C H S T D A  
2161 CAAGTTCCTT GCCGACGGCG GGTGCTCGGG GGGCGCTTAT GACATAATAA TTTGTGACGA GTGCCACTCC ACGGATGCCA  
GTTCAAGGAA CGGCTGCCGC CCACGAGCCC CCCGGAATA CTGTATTATT AAACACTGCT CACGGTGAGG TGCCTACGGT

+2 T S I L G I G T V L D Q A E T A G A R L V V L A T A T  
2241 CATCCATCTT GGCATTGGC ACTGTCTTG ACCAAGCAGA GACTGGGGG GCGAGACTGG TTGTGCTCGC CACCGCCACC  
GTAGGTAGAA CCCGTAACCG TGACAGGAAC TGGTTGCTCT CTGACGCCCC CGCTCTGACC AACACGAGCG GTGGCGGTGG

+2 P P G S V T V P H P N I E E V A L S T T G E I P F Y G  
2321 CCTCCGGGCT CCGTCACTGT CCCCCATCCC AACATCGAGG AGGTTGCTCT GTCCACCACC GGAGAGATCC CTTTTTACGG  
GGAGGCCCCA GGCAGTGACA CGGGGTAGGG TTGTAGCTCC TCCAAACGAGA CAGGTGGTGG CCTCTCTAGG GAAAAATGCC

+2 K A I P L E V I K G G R H L I F C H S K K K C D E L  
2401 CAAGGCTATC CCCCTCGAAG TAATCAAGGG GGGGAGACAT CTCATCTTCT GTCAATCAAA GAAGAAGTGC GACGAACCTCG  
GTTCCGATAG GGGGAGCTTC ATTAGTTCCC CCCCTCTCTA GAGTAGAAGA CAGTAAGTTT CTTCTTCACG CTGCTTGAGC

+2 A A K L V A L G I N A V A Y Y R G L D V S V I P T S G  
2481 CCGCAAAGCT GGTCCGATTG GGCATCAATG CCGTGGCCTA CTACCGGGGT CTTGACGTGT CCGTCATCCC GACCAGCGGC  
GGCGTTTTCGA CCAGCGTAAC CCGTAGTTAC GGCACCGGAT GATGGGGCCA GAACTGCACA GGCAGTAGGG CTGTCGCGCG

FIG. 5-4



pCMV-delNS35

+2 D V V V V A T D A L M T G Y T G D F D S V I D C N T C  
2561 GATGTTGTCG TCGTGGCAAC CGATGCCCTC ATGACCGGCT ATACCGGCGA CTTGCACTCG GTGATAGACT GCAATACGTC  
CTACAACAGC AGCACCGTTG GCTACGGGAG TACTGGCCGA TATGGCCGCT GAAGCTGAGC CACTATCTGA CGTTATGCAC

+2 V T Q T V D F S L D P T F T I E T I T L P Q D A V S  
2641 TGTCAACCCAG ACAGTCGATT TCAGCCTTGA CCCTACCTTC ACCATTGAGA CAATCAGGCT CCCCCAAGAT GCTGTCTCCC  
ACAGTGGGTC TGTACGCTAA AGTCGGAAC TGGATGGAAG TGTAACTCT GTTAGTGCGA GGGGGTTCTA CGACAGAGGG

+2 R T Q R R G R T G R G K P G I Y R F V A P G E R P S G  
2721 GCACTCAACG TCGGGGCAGG ACTGGCAGG GGAAGCCAGG CATCTACAGA TTTGTGGCAC CGGGGGAGCG CCCCTCCGGC  
CGTGAGTTGC AGCCCCGTC TGACCGTCCC CCTTGGGTCC GTAGATGTCT AAACACCGTG GCCCCCTCGC GGGGAGGCGC

+2 M F D S S V L C E C Y D A G C A W Y E L T P A E T T V  
2801 ATGTTGCACT CGTCCGTCCT CTGTGAGTGC TATGAGCCAG GCTCTGCTTG GTATGAGCTC AGCCCGCGCG AGACTACAGT  
TACAAGCTGA GCAGGCAGGA GACACTCAGG ATACTGCGTC CGACACGAAC CATACTCGAG TCGGGGCGGC TCTGATGTCA

+2 R L R A Y M N T P G L P V C Q D H L E F W E G V F T  
StuI

2881 TAGGCTACGA GCGTACATGA ACACCCCGGG GCTTCCCGTG TGCCAGGACC ATCTTGAATT TTGGGAGGGC GTCTTTACAG  
ATCCGATGCT CGCATGTACT TGTGGGGCCC CGAAGGGCAC ACGTCTCTGG TAGAACTTAA AACCTCCCCG CAGAAATGTC

+2 G L T H I D A H F L S Q T K Q S G E N L P Y L V A Y Q  
StuI  
----

2961 GCCTCACTCA TATAGATGCC CACTTTCTAT CCCAGACAAA GCAGAGTGGG GAGAACCTTC CTTACCTGGT AGCGTACCAA  
CGGAGTGAGT ATATCTACGG GTGAAAGATA GGGTCTGTTT CGTCTCACCC CTCTTGAAG GAATGGACCA TCGCATGGTT

FIG. 5-5





pCMV-delINS35

+2 A T V C A R A Q A P P P S W D Q M W K C L I R L K P T  
3041 GCCACCGTGT GCGCTAGGGC TCAAGCCCCT CCCCATCGT GGGACCATGT TTGATTGGCC TCAAGCCCAC  
CGGTGGCACA CGGATCCCG AGTTGGGGA GGGGTAGCA CCTGGTCTA CACCTTACA AACTAAGCGG AGTTGGGTG

---

+2 L H G P T P L L Y R L G A V Q N E I T L T H P V T K  
3121 CCTCCATGGG CCAACACCCC TGCTATACAG ACTGGGGCT GTTCAGATG AAATCACCTT GACGCACCCA GTCACCAAAT  
GGAGGTACCC GGTGTGGG ACGATATGTC TGACCCGCGA CAAGTCTTAC TTATGTGGA CTGGCTGGGT CAGTGTGTTA

---

+2 Y I M T C M S A D L E V V T S T W V L V G G V L A A L  
3201 ACATCATGAC ATGCATGTCG GCCGACCTGG AGTCTGTAC GAGCACCTGG GTGCTCGTTG GCGGCGTCCT GGCTGCTTTG  
TGTAAGTACTG TACGTACAGC CGGCTGGACC TCCAGCAGTG CTCGTGGACC CACGAGCAAC CGCCGCCAGGA CCGACGAAAC

---

+2 A A Y C L S T G C V V I V G R V V L S G K P A I I P D  
3281 GCGCGTATT GCCTGTCAAC AGGCTGCGTG GTCATAGTGG GCAGGGTCTG CTGTCCGGG AAGCCGGCAA TCATACCTGA  
CGGCGCATAA CGGACAGTTG TCCGACGCAC CAGTATCAC CGTCCAGCA GAACAGGCC TTCCGGCCGT AGTATGGACT

---

+2 R E V L Y R E F D E M E E C S Q H L P Y I E Q G M M  
3361 CAGGGAAGTC CTCTACCGAG AGTTGGATGA GATGGAAGAG TGCTCTCAGC ACTTACCGTA CATCGAGCAA GGGATGATGC  
GTCCCTTCAG GAGATGGCTC TCAAGCTACT CTACCTTCTC ACGAGAGTCG TGAATGGCAT GTAGCTCGTT CCTACTACG

---

+2 L A E Q F K Q K A L G L L Q T A S R Q A E V I A P A V  
3441 TCGCCGAGCA GTTCAAGCAG AAGGCCCTCG GCCTCCTGCA GACCGCGTCC CGTCAGGCAG AGTTATCGC CCCTGCTGTC  
AGCGGCTCGT CAAGTTCGTC TTCCGGGAGC CGGAGGACGT CTGGCGCAGG GCAGTCCGTC TCCAATAGCG GGGACGACAG

---

+2 Q T N W Q K L E T F W A K H M W N F I S G I Q Y L A G  
3521 CAGACCAACT GGCAGAAACT CGAGACCTTC TGGGCGAAGC ATATGTGAA CTTCATCAGT GGGATACAAT ACTTGGCGGG  
GTCTGGTTGA CCGTTTTTGA GCTCTGGAAG ACCCGCTTCG TATACACCTT GAAGTAGTCA CCTATGTTA TGAACCGCCC

FIG. 5-6



pCMV-delINS35

+2 L S T L P G N P A I A S L M A F T A A V T S P L T T  
3601 CTTGTCAACG CTGCTGGTA ACCCGCCCAT TGCCTCATTTG ATGGCTTTTA CAGCTGCTGT CACCAGCCCA CTAACCACTA  
GAACAGTTGC GACGGACCAT TGGGGCGGTA ACAGAACTAAC TACCGAAAT GTGGTGGGT GATTGGTGAT

+2 S Q T L L F N I L G G W V A A Q L A A P G A A T A F V  
3681 GCCAAACCCCT CCTCTTCAAC ATATTGGGG GGTGGTGGC TGCCAGCTC GCGCCCCCG GTCCGCTAC TGCCTTTGTG  
CGTTTGGGA GGAGAGTTG TATAACCCCC CCACCCACCG ACGGTCTGAG CCGCGGGGGC CACGGCGATG ACGGAAACAC

+2 G A G L A G A A I G S V G L G K V L I D I L A G Y G A  
3761 GCGGTGGCT TAGTGGCG CGCCATCGC AGTGTGGAC TGGGAAGGT CCTCATAGAC ATCTTGCG GGTATGGCG  
CGGGACCGA ATCGACCGG GCGGTAGCG TCACAACCTG ACCCTTCCA GGAGTATCTG TAGGAACGTC CCATACCGG

+2 G V A G A L V A F K I M S G E V P S T E D L V N L L  
3841 GGGCTGGCG GGAGCTCTTG TGGCATTCAG GATCATGAGC GGTGAGTCC CCTCCACGGA GGACCTGGT AATCTACTGC  
CCGACCGC CCTCGAGAAC ACCGTAAGT CTAGTACTCG CCACTCCAGG GGAGGTGCTT CCTGGACCAG TTAGATGACG

+2 P A I L S P G A L V V G V V C A A I L R R H V G P G E  
3921 CCGCATCCT CTGCGCCGGA GCGCTCGTAG TCGGCGTGGT CTGTGCAGCA ATACTGCGCC GGCACGTTGG CCGGGCGGAG  
GGCGTAGGA GAGCGGCGCT CGGGAGCATC AGCGGCACCA GACACGTCGT TATGACGCGG CCGTGCAACC GGGCCGCTC

+2 G A V Q W M N R L I A F A S R G N H V S P T H Y V P E  
4001 GGGCAGTGC AGTGGATGAA CCGGTGATA GCCTTGGCT CCGGGGGGAA CCATGTTTCC CCCACGCACT ACGTGCCGGA  
CCCCGTACG TCACCTACTT GGCGACTAT CGGAAGCGGA GGGCCCCCTT GGTACAAAGG GGTGCGTGA TGCACGGCT

+2 S D A A A R V T A I L S S L T V T Q L L R R L H Q W  
4081 GAGCATGCA GGTGCCGCG TCACCTGCCAT ACTCAGCAGC CTCAGTGTA CCCAGTCTCT GAGGCGACTG CACCACTGGA  
CTCGTACGT CGACGGGCG AGTGACGGTA TGAGTCTGCG GAGTGACATT GGTGCGAGGA CTCGCTGAC GTGTCACCT

FIG. 5-7



pCMV-delINS35

+2 I S S E C T T P C S G S W L R D I W D W I C E V L S D  
4161 TAAGCTCGGA GTGTACCACT CCATGCTCG GTTCCTGGCT AAGGGACATC TGGGACTGGA TATGCGAGGT GTTGAGCGAC  
ATTGAGCCT CACATGGTGA GGTACGAGGC CAAGGACCGA TTCCCTGTAG ACCCTGACCT ATACGCTCCA CAACTCGCTG

+2 F K T W L K A K L M P Q L P G I P F V S C Q R G Y K G

BamHI  
-----

4241 TTTAAGACCT GGCTAAAAGC TAAGCTCATG CCACAGCTGC CTGGGATCCC CTTTGTGTCC TGCCAGCGCG GGTATAAGGG  
AAATTCTGGA CCGATTTTCG ATTCGAGTAC GGTCTGACG GACCCTAGGG GAAACACAGG ACGGTGCGCG CCATATTCCC

+2 V W R G D G I M H T R C H C G A E I T G H V K N G T  
4321 GGTCTGGCGA GGGGACGGCA TCATGCCAC TCGCTGCCAC TGTGGAGCTG AGATCACTGG ACATGTCAA AACGGGACGA  
CCAGACCGCT CCCCTGCCGT AGTACGTGTG ACGACGGTG ACACCTCGAC TCTAGTGACC TGTACAGTTT TTGCCCTGCT

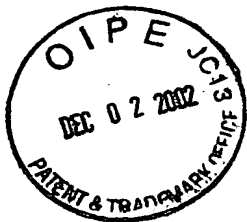
+2 M R I V G P R T C R N M W S G T F P I N A Y T T G P C  
4401 TCAGGATCGT CGGTCTTAGG ACCTGCAGGA ACATGTGGAG TGGGACCTTC CCCATTAATG CCTACACCAC GGGCCCCCTGT  
ACTCCTAGCA GCCAGGATCC TGGACGTCT TGTACACCTC ACCCTGGAAG GGTAATTAC GGATGTGGTG CCGGGGGACA

+2 T P L P A P N Y T F A L W R V S A E E Y V E I R Q V G  
4481 ACCCCCCTTC CTGCGCCGAA CTACACGTTG GCGCTATGGA GGTGTCTGC AGAGGAATAC GTGGAGATAA GGCAGGTGGG  
TGGGGGGAAG GACGGGGCTT GATGTGCAAG CGGATACCT CCCACAGACG TCTCCTTATG CACCTCTATT CCGTCCACCC

+2 D F H Y V T G M T T D N L K C P C Q V P S P E F F T  
4561 GGACTTCCAC TACGTGACGG GTATGACTAC TGACAATCTT AAATGCCGT GCCAGGTCCC ATCGCCCGAA TTTTTCACAG  
CCTGAAGGTG ATGCACTGCC CATACTGATG ACTGTTAGAA TTACGGGCA CGGTCCAGG TAGCGGGCTT AAAAAGTGC

+2 E L D G V R L H R F A P P C K P L L R E E V S F R V G  
4641 AATTGGACGG GGTGCGCCTA CATAGGTTTG CGCCCCCTG CAAGCCCTTG CTGCGGGAGG AGGTATCATT CAGAGTAGGA  
TTAACCTGCC CCACGGGGAT GTATCCAAAC CGGGGGGGAC GTTCGGGGAAC GACGCCCTCC TCCATAGTAA GTCTCATCCT

FIG. 5-8



pCMV-deINS35

+2 L H E Y P V G S Q L P C E P E P D V A V L T S M L T D  
4721 CTCCACGAAT ACCGGTAGG GTCGCAATTA CCTTGGAGC CCGAACCGGA GTGGCCGTG TTGACGTCCA TGCTCACTGA  
GAGGTGCTTA TGGGCCATCC CAGCGTTAAT GGAACGCTCG GGCTTGGCCT GCACCGGCAC AACTGCAGGT ACGAGTGACT

+2 P S H I T A E A A G R R L A R G S P P S V A S S S A  
4801 TCCCTCCCAT ATAACAGCAG AGCGGGCCGG GCGAAGGTG GCGAGGGGAT CACCCCCCTC TGTGGCCAGC TCCTCGGCTA  
AGGAGGGTA TATTGTCGTC TCCGCCGGCC CGCTTCCAAC CGCTCCCTA GTGGGGGAG ACACCGGTG AGGAGCCGAT

+2 S Q L S A P S L K A T C T A N H D S P D A E L I E A N  
4881 GCCAGCTATC CGCTCCATCT CTCAAGGCAA CTTCACCGC TAACCATGAC TCCCTGATG CTGAGCTCAT AGAGGCCAAC  
CGGTCGATAG GCGAGGTAGA GAGTCCGTT GAACGTGGCG ATTGTACTG AGGGGACTAC GACTCGAGTA TCTCCGGTTG

+2 L L W R Q E M G G N I T R V E S E N K V V I L D S F D  
4961 CTCCTATGGA GGCAGGAGAT GGGCGGCAAC ATCACCAGG TTGAGTCAGA AAACAAAGTG GTGATTCTGG ACTCCTTCCA  
GAGGATACCT CCGTCTCTA CCGCCGTTG TAGTGTTCC AACTCAGTCT TTTGTTTTCAC CACTAAGACC TGAGGAAGCT

+2 P L V A E E D E R E I S V P A E I L R K S R R F A Q  
5041 TCCGCTTGTG GCGGAGGAGG ACCAGCGGA GATCTCCGTA CCGCAGAAA TCCTGGGAA GTCTCGGAGA TTCGCCCAGG  
AGGCGAACAC CGCTCCTCC TGCTCGCCCT CTAGAGGCAT GGGCGTCTTT AGGACGCCTT CAGAGCCTCT AAGCGGTCC

+2 A L P V W A R P D Y N P P L V E T W K K P D Y E P P V  
5121 CCCTGCCCCGTTGGCGCGG CCGGACTATA ACCCCCGCT AGTGGAGACG TGGAAAAAGC CCGACTACGA ACCACCTGTG  
GGGACGGGCA AACCCGGGCG GGCCTGATAT TGGGGGCGA TCACCTCTGC ACCTTTTTCG GGCTGATGCT TGGTGGACAC

+2 V H G C P L P P P K S P P V P P P R K K R T V V L T E  
5201 GTCCATGGCT GCCCGCTTCC ACCTCCAAAG TCCCTCCTG TGGCTCGCC TCGGAAGAAG CGGACGGTGG TCCTCACTGA  
CAGGTACCGA CCGGCGAAGG TGGAGTTTC AGGGAGGAC ACGGAGGCG AGCTTCTTC GCCTGCCACC AGGAGTGACT

FIG. 5-9



pCMV-deINS35

+2 S T L S T A L A E L A T R S F G S S S T S G I T G D  
5281 ATCAACCCTA TCTACTGCCT TGGCCGAGCT CGCCACCAGA AGCTTTGGCA GCTCCTCAAC TTCCGGCATT ACGCGGGACA  
TAGTTGGGAT AGATGACGGA ACCGGCTCGA GCGGTGTCT TCGAAACCGT CGAGGAGTTG AAGGCCGTAA TGCCCGCTGT

+2 N T T S S E P A P S G C P P D S D A E S Y S S M P P  
5361 ATACGACAAC ATCCTCTGAG CCGGCCCTT CTGGCTGCC CCGGACTCC GACGTGAGT CCTATTCTC CATGCCCCCC  
TATGCTGTTG TAGGAGACTC GGGCGGGAA GACCGACGGG GGGGCTGAG CTGCGACTCA GGATAAGGAG GTACGGGGGG

+2 L E G E P G D P D L S D G S W S T V S S E A N A E D V  
BamHI

-----  
5441 CTGGAGGGG AGCCTCGCA TCCGGAATTT AGCGACGGT CATGTC AAC GGTCACTAGT GAGGCCAACG CGGAGGATGT  
GACCTCCCC TCGGACCCCT AGGCTAGAA TCGTGCCCA GTACCAGTTG CCAGTCATCA CTCCGGTTGC GCCTCCTACA

+2 V C C S M S Y S W T G A L V T P C A A E E Q K L P I  
5521 CGTGCTGC TCAATGCTT ACTCTGGAC AGGGCACTC GTACCCCGT GCGCCGCGA AGAACAGAAA CTGCCCATCA  
GCACACGACG AGTTACAGAA TGAGAACCTG TCCGGGTGAG CAGTGGGCA CGCGGGCCT TCTTGTCTT GACGGGTAGT

+2 N A L S N S L L R H H N L V Y S T T S R S A C Q R Q K  
5601 ATGCACTAAG CAATCTGTTG CTACGTCACC ACAATTTGGT GTATTCCACC ACCTCAGCA GTGCTTGCCA AAGCAGAAG  
TACGTGATTC GTTGAGCAAC GATGCAGTGG TGTAAACCA CATAAGTGG TGGAGTGGT CACGAACGGT TTCCGTCTTC

+2 K V T F D R L Q V L D S H Y Q D V L K E V K A A S K  
5681 AAAGTCACAT TTGACAGACT GCAATCTCTG GACAGCCATT ACCAGGACGT ACTCAAGGAG GTTAAAGCAG CGGCGTCAAA  
TTTTCAGTGA AACTGTCTGA CGTTCAAGAC CTGTCCGTAA TGGTCTGCA TGAGTTCCTC CAATTCTGTC GCGGCAGTTT

+2 V K A N L L S V E E A C S L T P P H S A K S K F G Y  
5761 AGTGAAGGCT AACTTGCTAT CCGTAGAGGA AGCTTGACG CTGACGCCCC CACACTCAGC CAAATCCAAG TTGTTATG  
TCACTTCCGA TTGAACGATA GGCATCTCCT TCGAACGTG GACTCGGGG GTGTGAGTCG GTTTAGGTTT AAACCAATAC

FIG. 5-10



pCMV-deINS35

+2 G A K D V R C H A R K A V T H I N S V W K D L L E D N  
5841 GGGCAAAAGA CGTCCGTTGC CATGCCAGAA AGGCCGTAAC CCACATCAAC TCCGTGTGGA AAGACCTTCT GGAAGACAAT  
CCCCGTTTCT GCAGGCAACG GTACGGTCTT TCCGGCAATG GGTGTAGTG AGGCACACCT TTCTGGAAGA CCTTCTGTIA

+2 V T P I D T T I M A K N E V F C V Q P E K G G R K P A  
5921 GTAACACCAA TAGACACTAC CATCATGGCT AAGAAGGAGG TTTTCTGGGT TCAGCCTGAG AAGGGGGTC GTAAGCCAGC  
CATTTGGTT ATCTGTGATG GTAGTACCGA TTCTTGCTCC AAAAGACGCA AGTCGGACTC TTCCCCCCCAG CATTCCGTCG

+2 R L I V F P D L G V R V C E K M A L Y D V V T K L P  
6001 TCGTCTCATC GTGTTCCCG ATCTGGGCGT GCGGTGTGC GAAAGATGG CTTTGTACGA CGTGGTTACA AAGCTCCCT  
AGCAGAGTAG CACAAGGGGC TAGACCCGCA GCGCACAGG CTTTCTTACC GAAACATGCT GCACCAATGT TTCGAGGGGA

+2 L A V M G S S Y G F Q Y S P G Q R V E F L V Q A W K S  
EcoRI  
-----

6081 TGGCCGTGAT GGAAGCTCC TACGGATTCC AATACTACC AGCACAGGG GTTGAATTC TCGTCAAGC GTGGAAGTCC  
ACGGCACTA CCCTTCGAGG ATGCCTAAGG TTATGAGTGG TCCTGTGCGC CAACTTAAG AGCACGTTCC CACCTTCAGG

+2 K K T P M G F S Y D T R C F D S T V T E S D I R T E E  
6161 AAGAAAACCC CAATGGGTT CTGATGAT ACCCGCTGCT TTGACTCCAC AGTCACTGAG AGCGACATCC GTACGGAGGA  
TTCTTTTGGG GTTACCCCAA GAGCATACTA TGGGGGACGA AACTGAGGTG TCAGTGACTC TCGCTGTAGG CATGCCCTCT

+2 A I Y Q C C D L D P Q A R V A I K S L T E R L Y V G  
6241 GGCAATCTAC CAATGTTGTG ACCTCGACCC CCAAGCCCGC GTGGCCATCA AGTCCCTCAC CGAGAGGCTT TATGTTGGGG  
CCGTTAGATG GTTACAACAC TGGAGCTGGG GGTTCGGCG CACCGGTAGT TCAGGAGATG GCTCTCCGAA ATACAACCCC

+2 G P L T N S R G E N C G Y R R C R A S G V L T T S C G  
6321 GCCCTCTTAC CAATTCAAGG GGGGAGAACT CCGGCTATCG CAGGTCCCGC GCGAGCGGG TACTGACAC TAGCTGTGT  
CGGGAGAATG GTTAAGTTCC CCCCCTCTGA CCGCGATAGC GTCCACGGCG CGCTCGCGC ATGACTGTG ATCGACACCA

FIG. 5-11



pCMV-delINS35

+2 N T L T C Y I K A R A A C R A A G L Q D C T M L V C G  
6401 AACACCCCTCA CTTGCTACAT CAAGGCCCGG GCAGCCTGTC GAGCCGAGG GCTCCAGGAC TGCACCATGC TCGTGTGTGG  
TTGTGGGAGT GAACGATGTA GTTCCGGGCC CGTCGGACAG CTCGGCGTCC CGAGGTCTTG AGTGTGACG AGCACACACC

---

+2 D D L V V I C E S A G V Q E D A A S L R A F T E A M  
6481 CGACGACTTA GTCGTTATCT GTGAAAGCGC GGGGTCCAG GAGGACGGG CGAGCCTGAG AGCTTCACG GAGGCTATGA  
GCTGCTGAAT CAGCAATAGA CACTTTCGGC CCCCAGGTC CTCCTGGCC GCTCGGACTC TCGGAAGTGC CTCGATACT

---

+2 T R Y S A P P G D P P Q P E Y D L E L I T S C S S N V  
6561 CCAGGTACTC CGCCCCCCT GGGACCCCC CACAACCAGA ATACGACTTG GAGCTCATAA CATCATGCTC CTCCAACGTG  
GGTCCATGAG GCGGGGGGA CCCCTGGGG GTGTGGTCT TATGCTGAAC CTCGAGTATT GTAGTACGAG GAGGTTGCAC

---

+2 S V A H D G A G K R V Y Y L T R D P T T P L A R A A W  
6641 TCAGTCGCCC ACGACGGCGG TGGAAAGAGG GTCTACTACC TCACCCGTGA CCCTACAACC CCCCTCGGA GAGTGCCTG  
AGTCAGCGG TGCTGCGCGG ACCTTCTCC CAGATGATGG AGTGGCACT GGGATGTTGG GGGAGCGCT CTCGAGCGC

---

+2 E T A R H T P V N S W L G N I I M F A P T L W A R M  
6721 GGAGACAGCA AGACACACTC CAGTCAATTC CTGGCTAGGC AACATAATCA TGTTGCCCC CACACTGTGG GCGAGGATGA  
CCTCTGTCGT TCTGTGTGAG GTCAGTTAAG GACCGATCCG TTGTATTAGT ACAACGGGG GTGTGACACC CGCTCCTACT

---

+2 I L M T H F F S V L I A R D Q L E Q A L D C E I Y G A  
6801 TACTGATGAC CCATTTCCTT AGCGTCCTTA TAGCCAGGGA CCAGCTTGAA CAGGCCCTCG ATTGCGACAT CTACGGGGCC  
ATGACTACTG GGTAAGAAA TCGCAGGAAT ATCGGTCCCT GGTGAACTT GTCCGGGAGC TAACGCTCTA GATGCCCGG

---

+2 C Y S I E P L D L P P I I Q R L H G L S A F S L H S Y  
6881 TGCTACTCCA TAGAACCCT GGATCTACCT CCAATCATTC AAAGACTCCA TGGCCTCAGC GCATTTTCAC TCCACAGTTA  
ACGATGAGT ATCTTGGTGA CCTAGATGGA GGTTAGTAAG TTTCTGAGT ACGGAGTGC CGTAAAAGTG AGTGTCAAT

FIG. 5-12



pCMV-delINS35

+2 S P G E I N R V A A C L R K L G V P P L R A W R H R  
6961 CTCTCCAGGT GAAATCAATA GGTGGCCGC ATGCCTCAGA AAACCTGGGG TACGCCCTT GCGAGCTTGG AGACACCGGG  
GAGAGTCCA CTTTAGTTAT CCCACCGCG TACGGAGTCT TTTGAACCCC ATGGCGGGAA CGCTCGAACC TCTGTGCCCC

+2 A R S V R A R L L A R G G R A A I C G K Y L F N W A V  
7041 CCCGAGCGT CCGGCTAGG CTTCTGGCCA GAGGAGCAG GGTGCCATA TGTGGCAAGT ACCTCTTCAA CTGGGCAGTA  
GGGCTCGCA GCGCGATCC GAAGACCGT CTCTCCGTC CCGACGGTAT ACACCGTTCA TGGAGAAGTT GACCCGTCAT

+2 R T K L K L T P I A A A G Q L D L S G W F T A G Y S G  
7121 AGAACAAAGC TCAAACTCAC TCCAATAGCG GCGGCTGCC AGCTGGACTT GTCCGGCTGG TTCACGGCTG GCTACAGCGG  
TCTGTTCG AGTTGAGTG AGTTATCG CCGGACCGG TCGACCTGAA CAGGCCGACC AAGTCCGAC CGATGTGCCC

+2 G D I Y H S V S H A R P R W I W F C L L L L A A G V  
7201 GGGAGACATT TATCACAGCG TGTCTCATGC CCGGCCCCG TGGATCTGGT TTTGGCTTACT CCTGCTTGCT GCAGGGGTAG  
CCCTCTGTAA ATAGTCTCG ACAGAGTACG GGCCGGGGG ACCTAGACCA AAACGGATGA GGACGAACGA CGTCCCCATC

+2 G I Y L L P N R  
7281 GCATCTACCT CCTCCCCAAC CGATGAAGGT TGGGGTAAAC ACTCCGGCCT AAAAAAAAAA AAAATCTAG AAAGCGCGCG  
CGTAGATGGA GGAGGGGTG CTTACTTCCA ACCCATTTG TGAGGCCGGA TTTTTTTTTT TTTTITAGATC TTTCCGCGCG

BamHI MluI

7361 CAAGATATCA AGGATCCACT ACGGTTAGA GCTCGCTGAT CAGGCTCGAC TGTGCTTCT AGTTGCCAGC CATCTGTTGT  
GTTCTATAGT TCCTAGGTGA TCGCAATCT CGAGCGACTA GTCGGAGCTG ACACGGAAGA TCAACGGTCTG GTAGACAACA

7441 TTGCCCCCTCC CCCGTGCCTT CCTTGACCCT GGAAGGTGCC ACTCCCACTG TCCTTTCCCTA ATAAATGAG GAAATTGCAT  
AACGGGAGG GGGCACGGAA GGAAGTGGGA CCTTCCACGG TGAGGGTGAC AGGAAAGGAT TATTTTACTC CTTTAACGTA

FIG. 5-13



pCMV-deINS35

7521	CGCATTGTCT GAGTAGGTGT CATTCATTTC TGGGGGGTGG GGTGGGGCAG GACAGCAAGG GGGAGGATTG GGAAGACAAT GGGTAACAGA CTCATCCACA GTAAGATAAG ACCCCCCACC CCACCCCGTC CTGTGTTTCC CCCTCCTAAC CCTTCTGTTA
7601	AGCAGGCATG CTGGGGAGCT CTTCCGGTTC CTGGCTCACT GACTGGCTGC GCTGGTCTGT TCGGCTGCGG CGAGCGGTAT TCGTCCGTAC GACCCCTCGA GAAAGGCAAG GAGCGAGTGA CTGAGCGACG CGAGCCAGCA AGCCGACGCC GCTCGCCATA
7681	CAGCTCACTC AAAGGCGGTA ATAGGGTTAT CCACAGATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG GTCGAGTGAG TTTCCGCCAT TATGCCAATA GGTGTCTTAG TCCCCATTG CGTCCCTTCT TGTACACTCG TTTTCCGGTC
7761	CAAAAGGCCA GGAACCGTAA AAAGGCGCG TTTTCCATAG GCTCCGCCCC CCTGACGAGC ATCACAAAAA GTTTTCGGT CCTTGGCATT TTTCCGGCG AACGACCGCA AAAAGGTATC CGAGGCGGGG GGAAGTCTCG TAGTGTTTT
7841	TCGACGCTCA AGTCAGAGGT GCGGAAACCC GACAGGACTA TAAAGATACC AGGCGTTTCC CCCTGGAAGC TCCCTCGTGC AGCTGCGAGT TCAGTCTCCA CCGCTTTGGG CTGTCTTGAT ATTTCTATGG TCCGCAAGG GGGACCTTCG AGGAGCACG
7921	GCTCTCCTGT TCCGACCCTG CCGCTTACCG GATACCTGTC CGCTTTCTC CCTTCGGGA GCGTGGCGCT TTCTCAATGC CGAGAGGACA AGGCTGGGAC GCGCAATGGC CTATGGACAG GCGGAAAGAG GGAAGCCCTT CGCACCCGCGA AAGAGTTACG
8001	TCACGCTGTA GGTATCTCAG TTCGGTGTAG GTCGTTGCT CCAAGCTGGG CTGTGTGCAC GAACCCCGCG TTCAGCCCGA AGTCCGACAT CCATAGAGTC AAGCCACATC CAGCAAGCGA GTTCGACCC GACACACGTG CTTGGGGGGC AAGTGGGGCT
8081	CCGCTGCGCC TTATCCGGTA ACTATCGTCT TCACTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG GGCGACCGCG AATAGGCCAT TGATAGCAGA ACTCAGGTTG GGCCATTCTG TGCTGAATAG CCGTGACCGT CGTCGGTGAC
8161	GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG AAGTGGTGGC CTAACACTAG CTACACTAGA CATTGTCTTA ATCGTCTCGC TCCATACATC CGCCACGATG TCTCAAGAAC TTCACCACCG GATTGATGCC GATGTGATCT

FIG. 5-14



pCMV-deINS35

8241 AGGACAGTAT TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAAAACA  
TCCTGTGATA AACCATAGAC GCGAGACGAC TTGGGTCAAT GGAAGCCTTT TTCTCAACCA TCGAGAACTA GGCCGTTTGT

---

8321 AACCACCGCT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT  
TTGGTGGCGA CCATCGCCAC CAAAAAACA AACGTTCTGC GTCTAATGGC CGTCCTTTTT TCCTAGAGTT CTCTAGGAA

---

8401 TGATCTTTTC TACGGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA GGGATTTTGG TCATGAGATT ATCAAAAAGG  
ACTAGAAAAG ATGCCCCAGA CTGCGAGTCA CCTTGCTTTT GAGTGCAATT CCCTAAAACC AGTACTCTAA TAGTTTTTCC

---

8481 ATCTTCACCT AGATCCTTTT AAATTAAAA TGAAGTTTTA AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG  
TAGAAGTGA TCTAGGAAA TTTAATTTT ACTTCAAAAT TTAGTTAGAT TTCATATATA CTCATTTGAA CCAGACTGTC

---

8561 TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGCTATTTC GTTCATCCAT AGTTGCCTGA CTCCCCGTCG  
AATGTTACG AATTAGTCAC TCCGTGGATA GAGTCGCTAG ACAGATAAAG CAAGTAGGTA TCAACGGACT GAGGGGCAGC

---

8641 TGTAGATAAC TACGATACGG GAGGGCTTAC CATCTGGCC CAGTCTGCA ATGATACCGC GAGACCCACG CTCACCGGCT  
ACATCTATTG ATGCTATGCC CTCCGGAATG GTAGACGGG GTACACGACGT TACTATGGC CTCTGGGTGC GAGTGGCCGA

---

8721 CCAGATTTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG AGCGCAGAAG TGGTCCCTGCA ACTTTATCCG CCTCCATCCA  
GGTCTAAATA GTCGTTATT GGTGGGTCGG CCTTCCCGGC TCGGCTCTC ACCAGGACGT TGAATAGGC GGAGGTAGGT

---

8801 GTCTATTAAAT TGTGCCCCG AAGCTAGAGT AAGTAGTTTG CCAGTTAATA GTTTGGGCAA CGTTGTTGCC ATTGCTACAG  
CAGATAATTA ACAACGGCCC TTCGATCTCA TTCATCAAGC GGTCAATTAT CAAACGGCTT GCAACAACGG TAACGATGTC

---

8881 GCATCGTGGT GTCACGCTCG TCGTTTGGTA TGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGGCGAGT TACATGATCC  
CGTAGCACCA CAGTGGGAGC AGCAAAACCAT ACCGAAAGTAA GTCGAGGCCA AGGTTTGCTA GTTCCGCTCA ATGTACTAGG

FIG. 5-15



pCMV-deINS35

8961	CCCATGTTGT	GCAAAAAAGC	GGTAGCTCC	TTCCGTCCTC	CGATCGTTGT	CAGAACTAAG	TTGGCCGCAG	TGTTATCACT
	GGGTACAACA	CGTTTTTTTCG	CCAATCGAGG	AAGCCAGGAG	GCTAGCAACA	GTCTTCATTTC	AACCGGCGTC	ACAAATAGTGA
9041	CATGGTTATG	GCAGCACTGC	ATAATTCTCT	TACTGTCTATG	CCATCCGTAA	GATGCTTTTC	TGTGACTGGT	GAGTACTCAA
	GTACCAATAC	CGTCGTGACG	TATTAAGAGA	ATGACAGTAC	GGTAGGCATT	CTACGAAAAG	ACACTGACCA	CTCATGAGTT
9121	CCAAGTCATT	CTGAGAAATAG	TGTATGCGGC	GACCGAGTTG	CTCTTGCCCG	GGCTCAATAC	GGGATAATAC	CGCGCCACAT
	GGTTCAGTAA	GACTCTTATC	ACATACGCCG	CTGGCTCAAC	GAGAACGGGC	CGCAGTTATG	CCCTATTATG	GCGCGGTGTA
9201	AGCAGAACTT	TAAAAGTGCT	CATCATTGGA	AAACGTTCTT	CGGGCGGAAA	ACTCTCAAGG	ATCTTACCGC	TGTTGAGATC
	TCGTCTTGAA	ATTTTCACGA	GTAATAACCT	TTTGCAAGAA	GCCCCGCTTT	TGAGAGTTCC	TAGAAATGGC	ACAACTCTAG
9281	CAGTTCGATG	TAACCCACTC	GTGCACCCAA	CTGATCTTCA	GCATCTTTTA	CTTTCAACCAG	CGTTTCTGGG	TGAGCAAAAA
	GTCAAGCTAC	ATTGGGTGAG	CACGTGGGTT	GACTAGAACT	CGTAGAAAAT	GAAAGTGGTC	GCAAAGACCC	ACTCGTTTTT
9361	CAGGAAGGCA	AAATGCCGCA	AAAAAGGAA	TAAGGGCGAC	ACGGAATGT	TGAATACTCA	TACTCTTCCT	TTTTCAATAT
	GTCCCTTCCGT	TTTACGGCGT	TTTTTCCCTT	ATTCCCGCTG	TGCCTTTACA	ACTTATGAGT	ATGAGAAGGA	AAAAGTTATA
9441	TATTGAAGCA	TTTATCAGGG	TTATTGTCTC	ATGAGCGGAT	ACATATTGA	ATGTATTAG	AAAAATAAAC	AAATAGGGGT
	ATAACTTCGT	AAATAGTCCC	AATAACAGAG	TACTCGCCTA	TGTATAAACT	TACATAAATC	TTTTTTATTG	TTTATCCCCA
9521	TCCGGGCACA	TTTCCCCGAA	AAGTCCACC	TGACGTCTAA	GAACCATTA	TTATCATGAC	ATTAACCTAT	AAAAATAGGC
	AGCGCGGTGT	AAAGGGGCTT	TTACGGGTGG	ACTGCAGATT	CTTTGGTAAT	AATAGTACTG	TAATTGGATA	TTTTTATCCG
9601	GTATCACGAG	GCCCTTTCTG	C					
	CATAGTGCTC	CGGAAAGCA	G					

FIG. 5-16



pCMV-II

1	TCGCGCGGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGCT GTAAAGCGGAT AGCGCGCAAA GCCACTACTG CCACCTTTGG AGACTGTGTA CGTCGAGGGC CTCTGCCAGT GTCGAACAGA CATTCGGCTA
81	GCGGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGGGGGTG TTGGCGGGTG TCGGGGGCTGG CTTAACATATG CGGCATCAGA CGGCCCTCGT CTGTTCCGGC AGTCCCGCGC AGTCGCCAC AACGCCCCAC AGCCCGGACC GAATTGATAC GCCGTAGTCT
161	GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTGTCA AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG CGTCTAACAT GACTCTCAG TGGTATACTT CGAAAAACGT TTTCGGATCC GGAGGTTTT TCGGAGGAGT GATGAAGACC
241	AATAGCTCAG AGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAT TAGTCAGCCA TGGGGCGGAG AATGGGCGGA TTATCGAGTC TCCGGCTCCG CCGGAGCCGG AGACGTATT ATTITTTTIA ATCAGTCGT ACCCGGCTC TTACCGCGCT
321	ACTGGCGGG GAGGAATTA TTGGCTATTG GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT TGACCGGCCC CTCCTTAAT AACCGATAAC CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA
401	CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT ATCAATTAGT AATGCCCCAG TAATCAAGTA
481	AGCCCATATA TGGAGTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCGGCCCAT TCGGGTATAT ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GGGCGGGTAA
561	GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTACGGT CTGCAGTTAT TACTGCATAC AAGGTATCA TTGGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA
641	AAACTGCCCC CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG TCAATGACGG TAAATGGCCC TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGGCGG GGATAACTGC AGTTACTGCC ATTTACCGGG

FIG. 7-1

pCMV-II

721 GCCTGGCATT ATGCCCAGTA CATGACCCTTA CGGGACCTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC  
CGGACCGTAA TACGGGTCTAT GTACTGGAAT GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

---

801 CATGGTGATG CGGTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA  
GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATCGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGGTGGGT

---

881 TTGACGTCAA TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCTGA ATAAACCCCG CCCGTTGACG  
AACTGCAGTT ACCCTCAAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG GGGCAACTGC

---

961 CAAATGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG  
GTTTACCCGC CATCCGCACA TGCCACCCCTC CAGATATATT CGTCTGAGC AAATCACTTG GCAGTCTAGC GGAACCTCTGC

---

1041 CCATCCACGC TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGGCG GGAACGGTGC ATTGGAACGC  
GGTAGGTGG ACAAACCTGG AGGTATCTTC TGTGGCCCTG GCTAGGTGG AGGCGCCGC CCTTGCCACG TAACCTTGGC

---

1121 GGATTCGCC TGCCAAGAGT GACGTAAGTA CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA  
CCTAAGGGGC ACGGTTCTCA CTGCATTCTAT GCGGATATC TGAGATATCC GTGTGGGGA ACCGAGAATA CGTACGATAT

---

1201 CTGTTTTTGG CTTGGGGCCT ATACACCCCC GCTCCTTATG CTATAGGTGA TGGTATAGCT TAGCCTATAG GTGTGGGTTA  
GACAAAACC GAACCCCGGA TATGTGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCCAAT

---

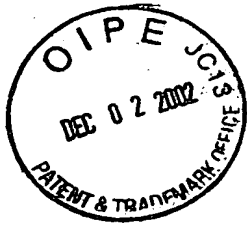
1281 TTGACCCTTA TTGACCCTC CCCTATTGGT GACGATACCT TCCATTACTA ATCCATAACA TGGCTCTTTG CCACAACTAT  
AACTGGTAAT AACTGGTGAG GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC GGTGTTGATA

---

1361 CTCATTTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTTAT  
GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCCTGA GACATAAAAA TGTCTTACCC CAGGTAAATA

---

FIG. 7-2



pCMV-II

```
1441 TATTTACAAA TTCACATATA CAACAACGCC GTCCCCCGTG CCGGAGTTT TTATTAAACA TAGCGTGGGA TCTCCGACAT
    ATAAATGTTT AAGTGTATAT GTTGTGCGG CAGGGGGCAC GGGCGTCAAA AATAATTGT ATCGCACCTT AGAGGCTGTA

1521 CTCGGGTACG TGTTCGGGAC ATGGGCTCTT CTCCGGTAGC GGGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA
    GAGCCCATGC ACAAGGCCCTG TACCCGAGAA GAGGCCATCG CCGCCTCGAA GGTGTAGGT CGGGACCAGG GTAGGCAGGT

1601 GCGGCTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA CTTAGGCACA GCACAATGCC CACCACCACC
    CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT GAATCCGTGT CGTGTACGG GTGGTGGTGG

1681 AGTGTCCCG ACAAGGCCGT GCGGTAGGG TATGTCTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT GGACGCAGAT
    TCACACGGCG TGTTCGGCA CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGGG CCTGCGTCTA

1761 GGAAGACTTA AGGCAGCGG AGAAGAAGAT GCAGGCAGCT GAGTTGTTGT ATTCTGATAA GAGTCAAGGG TAACTCCCGT
    CCTTCTGAAT TCCGTCCCGG TCTTCTTCTA CCTCCGTCGA CTCACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA

1841 TCGGGTGCTG TTAACGGTGG AGGCAGTGT AGTCTCAGCA GTACTCGTTG CTGCCGGCGG CGCCACCAGA CATAATAGCT
    ACGCCACGAC AATTGCCACC TCCCGTCACA TCAGACTGCT CATGAGCAAC GACGGCGCGG GCGGTGGTCT GTATTATCGA

1921 GACAGACTAA CAGACTGTTC CTTTCCATGG GTCFTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCAGA CTCGAGCAAG
    CTGTCTGATT GTCTGACAAG GAAAGGTACC CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTCT GAGCTCGTTC

2001 TCTAGAAAGG CGCGCCAAGA TATCAAGCAT CCACTACGGG TTAGAGTCTG CTGATCAGCC TCGACTGTGC CTTCTAGTTG
    AGATCTTCC GCGGGTTCT ATAGTTCTTA GGTGATGGC AATCTCGAG GACTAGTCGG AGCTGACAGG GAAGATCAAC
```

EcoRI

-----

BamHI MluI

XbaI

-----

FIG. 7-3



pCMV-II

2081 CCAGCCATCT GTTGTITGCC CCTCCCCCGT GCCTTCCTTG ACCCTGGAAG GTGCCACTCC CACTGTCTCTT TCCTAATAAA  
GCTCGGTAGA CAACAAACGG GGAGGGGGCA CGGAAGGAAC TGGGACCTTC CACGGTGAGG GTGACAGGAA AGGATTATTT

---

2161 ATGAGGAAAT TGCATCGCAT TGTCTGAGTA GGTGTCATTC TATTCTGGG GGTGGGTGG GGCAGGACAG CAAGGGGGAG  
TACTCCTTTA ACGTAGCGTA ACAGACTCAT CCACAGTAAG ATAAGACCCC CCACCCCACC CCGTCTCTGTG GTTCCCCCTC

---

2241 GATTGGGAAG ACAATAGCAG GCATGCTGGG GAGCTCTTCC GCTTCCTCGC TCACCTGACTC GCTGCGCTCG GTCTTCTGGC  
CTAACCCCTC TGTATCGTC CGTACGACCC CTCGAGAAGG CGAAGGAGCG AGTGACTGAG CGACGGGAGC CAGCAAGCCC

---

2321 TCGGGCGAGC GGTATCAGCT CACTCAAAGG CGGTAATACG GTTATCCACA GAATCAGGGG ATAACGCAGG AAAGAACATG  
ACGCCGCTCG CCATAGTCGA GTGAGTTTCC GCCATTATGC CAATAGGTGT CTTAGTCCCC TATTGCGTCC TTTCTTGTA

---

2401 TGAGCAAAAG GCCAGCAAAA GGCCAGGAAC CGTAAAAAGG CCGCGTTGCT GCGGTTTTTC CATAGGCTCC GCCCCCTGA  
ACTCGTTTTC CCGTCGTTTT CCGTCTCTTG GCATTTTTC GCGCAACGA CCGCAAAAAG GTATCCGAGG CGGGGGGACT

---

2481 CGAGCATCAC AAAAATCGAC GCTCAAGTCA GAGGTGGCGA AACCCGACAG GACTATAAAG ATACCAGGCG TTTCCCCCTG  
GCTCGTAGTG TTTTITAGTG CGAGTTCAGT CTCACCCGCT TTGGGCTGTC CTGATAATTTC TATGGTCCGC AAAGGGGAC

---

2561 GAAGTCCCT CGTGGGCTCT CCTGTTCGA CCTGCGGCT TACCGGATAC CTGTCCGCTT TTTCTCCCTC GGAAGCGTG  
CTTCGAGGGA GCACGGGAGA GAACAAGGCT GGGACGGCGA ATGGCTATG GACAGGCGGA AAGAGGGAAG CCTTTCGCAC

---

2641 GCGCTTCTC AATGCTCAG CTGTAGGTAT CTCAGTTCCG TGTAGGTGCT TCGTCCAAAG CTGGGCTGTG TGCACGAACC  
CGCGAAAGAG TTACGAGTGC GACATCCATA GAGTCAAGCC ACATCCAGCA AGCGAGTTTC GACCCGACAC ACGTGTCTGG

---

2721 CCCCCCTCAG CCCGACCGCT GCGCCTTATC CGGTAACATAT CGTCTTGAGT CCAACCCGCT AAGACACGAC TTATCGCCAC  
GGGGCAAGTC GGGCTGGCGA CGCGGAATAG GCCATTGATA GCAGAACTCA GGTGGGCA TTCTGTGCTG AATAGCGGTG

FIG. 7-4

pCMV-II

2801 TGGCAGCAGC CACTGGTAAC AGGATTAGCA GAGCGAGTA TGTAGCGGT GCTACAGAGT TCTTGAAGTG GTGGCCTAAC  
ACCGTCTCG GTGACCAATTG TCCTAATCGT CTGCTCCAT ACATCGGCA CGATGTCTCA AGAATTTCAC CACCGGATTG

---

2881 TAGGGCTACA CTAGAAAGGAC AGTATTGGT ATCTGGCTC TGCTGAAGCC AGTTACCTTC GGA AAAAGAG TTGGTAGCTC  
ATGCCGATGT GATCTTCTG TCATAAACCA TAGACGGGAG ACGACTTCGG TCAATGGAAG CCTTTTCTC AACCATCGAG

---

2961 TTGATCCGGC AAACAAACCA CCGCTGGTAG CGGTGGTTT TTGTTTGCA AGCAGCAGAT TACGGCGAGA AAAAAAGGAT  
AACTAGGCCG TTTGTTTGGT GCGGACCATC GCCACCAAAA AAACAAACGT TCGTCTCTA ATGCGGTCT TTTTTTCTA

---

3041 CTCAGAAGA TCCTTTGATC TTTTCTACGG GGTCTGACGC TCAGTGGAAC GAAACTCAC GTTAAGGGAT TTTGGTCATG  
GAGTTCTTCT AGGAACTAG AAAAGATGCC CCAGACTGCG AGTCACCTTG CTTTGAGTG CAATCCCTA AAACCACTAC

---

3121 AGATTATCAA AAAGGATCTT CACCTAGATC CTTTAAAT AAAAATGAAG TTTTAAATCA ATCTAAAGTA TATATGAGTA  
TCTAATAGTT TTTCTAGAA GTGGATCTAG GAAATTTAA TTTTACTTC AAAATTAGT TAGATTTTAT ATATACTCAT

---

3201 AACTTGGTCT GACAGTTACC AATGCTTAAT CAGTGAGGCA CCTATCTCAG CGATCTGTCT ATTTCTGTCA TCCATAGTTG  
TTGAACCCAGA CTGTCAATGG TTACGAATTA GTCACTCCGT GGATAGAGTC GCTAGACAGA TAAAGCAAGT AGGTATCAAC

---

3281 CCTGACTCCC CGTCGTGTAG ATAACTACGA TACGGGAGGG CTTACCATCT GCCCCCAGTG CTGCAATGAT ACCGCGAGAC  
GGACTGAGGG GCAGCACATC TATTGATGCT ATGCCCTCCC GAATGGTAGA CCGGGGTAC GACGTIAC TA TGGCGTCTG

---

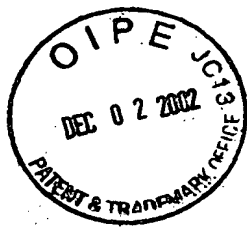
3361 CCACGCTCAC CGGTCCAGA TTTATCAGCA ATAAACCAGC CAGCCGGAAG GGCCGAGCGC AGAAGTGGT CTGCAACTTT  
GGTGGAGTG GCGAGGTCT AAATAGTCTG TATTGGTCTG GTCCGCTTC CCGGTCTCG TCTTCACCAG GACGTTGAAA

---

3441 ATCCGCCTCC ATCCAGTCTA TTAATTGTTG CCGGGAAGCT AGAGTAAGTA GTTCGCCAGT TAATAGTTTG CGCAACGTTG  
TAGGCGGAGG TAGTCAGAT AATTACAAC GGCCCTTGA TCTCATTCAT CAAGCGGTCA ATTATCAAAAC GCGTTGCAAC

FIG. 7-5

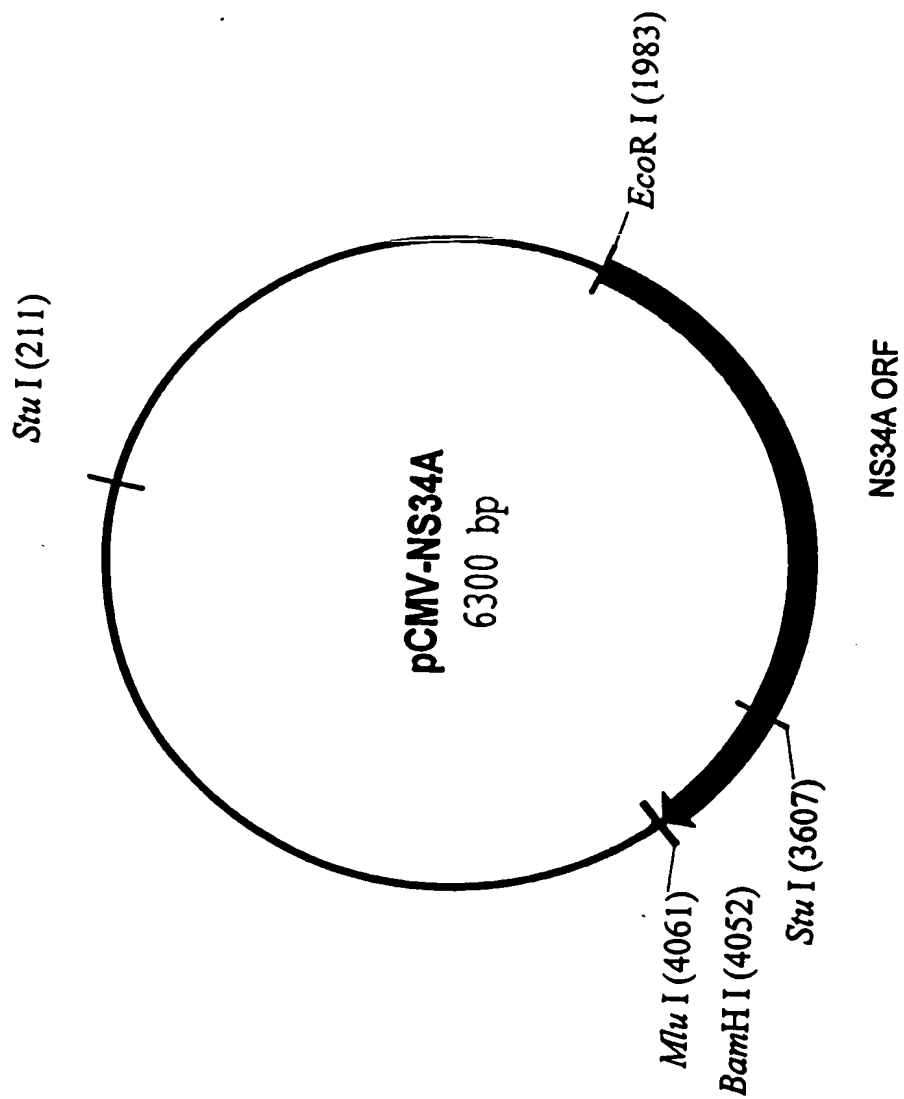
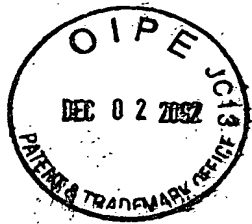




pCMV-II

3521	TTGCCATTGC TACAGGCATC GTGGTGTAC GCTCGTCGTT TGGTATGGCT TCATTCAGCT CCGGTTCCCA ACGATCAAGG AACGGTAACG ATGTCGGTAG CACCACAGTG CGAGCAGCAA ACCATACCGA AGTAAGTCGA GGCCAAGGCT TGCTAGTTCC
3601	CGAGTTACAT GATCCCCCAT GTTGTGCAAA AAAGCGGTTA GCTCCTTCGG TCCTCCGATC GTTGTGAGAA GTAAAGTTGGC GCTCAATGTA CTAGGGGGTA CAACACGTTT TTTCGCCAAT CGAGGAAGCC AGGAGGCTAG CAACAGTCTT CATTCACCCG
3681	CGCAGTGTTA TCACTCATGG TTATGGCAGC ACTGCATAAT TCTCTTACTG TCATGCCATC CGTAAGATGC TTTTCTGTGA GCGTCACAAT AGTGAGTACC AATACCGTCG TGACGTATTA AGAGAAATGAC AGTACGGTAG GCATTCTACG AAAAGACACT
3761	CTGGTGAGTA CTC AACCAAG TCATTCTGAG AATAGTGTAT CGGCGACCG AGTTGCTCTT GCCCGGGTC AATACGGGAT GACCACTCAT GAGTTGGTTC AGTAAGACTC TTATCACATA CGCGCTGGC TCAACGAGAA CGGGCCGCG TTTATGCCCTA
3841	AATACCGCGC CACATAGCAG AACTTTAAAA GTGCTCATCA TTGGAACCG TTCTTCGGGG CGAAAACTCT CAAGGATCTT TTATGGCGCG GTGTATCGTC TTGAAATTTT CACGAGTAGT AACCTTTTGC AAGAAGCCCC GCTTTTGAGA GTTCCTAGAA
3921	ACCGCTGTTG AGATCCAGTT CGATGTAACC CACTCGTGCA CCCAACTGAT CTTCAGCATC TTTTACTTTC ACCAGCGTTT TGGCGACAAC TCTAGGTCAA GCTACATTGG GTGAGCACGT GGGTTGACTA GAACTCGTAG AAAATGAAAG TGGTCGCAAA
4001	CTGGGTGAGC AAAAAACAGGA AGGCAAAATG CCGCAAAAAA GGAATAAGG GCGACACGGA AATGTTGAAT ACTCATACTC GACCCACTCG TTTTGTCTT TCCGTTTAC GCGTTTTT CCGTTATTCC CGCTGTGCCT TTACAACCTTA TGAGTATGAG
4081	TTCCTTTTTC AATATTATTG AAGCATTAT CAGGGTTATT GTCTCATGAG CGGATACATA TTTGAATGTA TTTAGAAAAA AAGGAAAAAG TTATAATAAC TTCGTAAATA GTCCCAATAA CAGAGTACTC GCCTATGTAT AAACCTTACAT AAATCTTTTT
4161	TAAACAAATA GGGGTTCGCG GCACATTTC CCGAAAAAGTG CCACCTGACG TCTAAGAAAC CATTATTATC ATGACATTAA ATTTGTTTAT CCCCAAGGCG CGTGTAAGG GCGTTTTTAC GGTGGACTGC AGATTCTTTG GTAATAATAG TACTGTAAAT
4241	CCTATAAAAA TAGCGGTATC ACGAGGCCCT TTCGTC GGATATTTT ATCCGCATAG TGCTCCGGGA AAGCAG

FIG. 7-6



**FIG. 8**



pCMV-NS34A

1 TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG  
AGCGCGCAAA GCCACTACTG CCACCTTTTGG AGACTGTGTA CGTCGAGGGC

---

51 GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG  
CTCTGCCAGT GTCGAACAGA CATTGCGCTA CGGCCCTCGT CTGTTGCGGC

---

101 TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG  
AGTCCCGCGC AGTCGCCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC

---

151 CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTTGCA  
GCCGTAGTGT CGTCTAACAT GACTCTCACG TGGTATACTT CGAAAAACGT

---

StuI  
~~~~~

201 AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTA CTCTTCTGG AATAGCTCAG  
TTTCGGATCC GGAGGTTTTT TCGGAGGAGT GATGAAGACC TTATCGAGTC

---

251 AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA  
TCCGGCTCCG CCGGAGCCGG AGACGTATTT ATTTTTTTTA ATCAGTCGGT

---

301 TGGGGCGGAG AATGGGCGGA ACTGGGCGGG GAGGGAATTA TTGGCTATTG  
ACCCGCGCTC TTACCCGCCT TGACCCGCCC CTCCCTTAAT AACCGATAAC

---

351 GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT  
CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

---

401 CATGTCCAAT ATGACCGCCA TGTGACATT GATTATTGAC TAGTTATTAA  
GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT

---

451 TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA TGGAGTTCGG  
ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT ACCTCAAGGC

---

501 CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC  
GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG

---

551 CCCGCCCAT TACGTCATAA ATGACGTATG TTCCCATAGT AACGCCAATA  
GGGCGGGTAA CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT

---

601 GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AAAGTACCGA  
CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA TTTGACGGGT

---

651 CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG  
GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC

---

701 TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCAGTA CATGACCTTA  
AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT GTACTGGAAT

---

751 CGGGACTTTC CTA CTCTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC  
GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

---

801 CATGGTGATG CGGTTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG  
GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATCGCCAAAC

---

851 ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG  
TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT ACCCTCAAAC

FIG. 9-1



pCMV-NS34A

|      |                           |                          |                          |                           |                           |
|------|---------------------------|--------------------------|--------------------------|---------------------------|---------------------------|
| 901  | TTTTGGCACC<br>AAAACCGTGG  | AAAATCAACG<br>TTTGTAGTTC | GGACTTTCCA<br>CCTGAAAGGT | AAATGTCGTA<br>TTTACAGCAT  | ATAACCCCGC<br>TATTGGGGCG  |
| 951  | CCCGTTGACG<br>GGGCAACTGC  | CAAATGGGCG<br>GTTTACCCGC | GTAGGCGTGT<br>CATCCGCACA | ACGGTGGGAG<br>TGCCACCCCTC | GTCTATATAA<br>CAGATATATT  |
| 1001 | GCAGAGCTCG<br>CGTCTCGAGC  | TTTAGTGAAC<br>AAATCACTTG | CGTCAGATCG<br>GCAGTCTAGC | CCTGGAGACG<br>GGACCTCTGC  | CCATCCACGC<br>GGTAGGTGCG  |
| 1051 | TGTTTTGACC<br>ACAAAACCTGG | TCCATAGAAG<br>AGGTATCTTC | ACACCGGGAC<br>TGTGGCCCTG | CGATCCAGCC<br>GCTAGGTCGG  | TCCGCGGCCG<br>AGGCGCCGGC  |
| 1101 | GGAACGGTGC<br>CCTTGCCACG  | ATTGGAACGC<br>TAACCTTGCG | GGATTCCCCG<br>CCTAAGGGGC | TGCCAAGAGT<br>ACGGTTCTCA  | GACGTAAGTA<br>CTGCATTTCAT |
| 1151 | CCGCCTATAG<br>GGCGGATATC  | ACTCTATAGG<br>TGAGATATCC | CACACCCCTT<br>GTGTGGGGAA | TGGCTCTTAT<br>ACCGAGAATA  | GCATGCTATA<br>CGTACGATAT  |
| 1201 | CTGTTTTTGG<br>GACAAAAACC  | CTTGGGGCCT<br>GAACCCCGGA | ATACACCCCC<br>TATGTGGGGG | GCTCCTTATG<br>CGAGGAATAC  | CTATAGGTGA<br>GATATCCACT  |
| 1251 | TGGTATAGCT<br>ACCATATCGA  | TAGCCTATAG<br>ATCGGATATC | GTGTGGGTTA<br>CACACCCAAT | TTGACCATTA<br>AACTGGTAAT  | TTGACCACTC<br>AACTGGTGAG  |
| 1301 | CCCTATTGGT<br>GGGATAACCA  | GACGATACTT<br>CTGCTATGAA | TCCATTACTA<br>AGGTAATGAT | ATCCATAACA<br>TAGGTATTGT  | TGGCTCTTTG<br>ACCGAGAAAC  |
| 1351 | CCACAACAT<br>GGTGTGATA    | CTCTATTGGC<br>GAGATAACCG | TATATGCCAA<br>ATATACGGTT | TACTCTGTCC<br>ATGAGACAGG  | TTCAGAGACT<br>AAGTCTCTGA  |
| 1401 | GACACGGACT<br>CTGTGCCTGA  | CTGTATTTTT<br>GACATAAAAA | ACAGGATGGG<br>TGTCCTACCC | GTCCATTTAT<br>CAGGTAAATA  | TATTTACAAA<br>ATAAATGTTT  |
| 1451 | TTCACATATA<br>AAGTGTATAT  | CAACAACGCC<br>GTTGTTGCCG | GTCCCCCGTG<br>CAGGGGGCAC | CCCGCAGTTT<br>GGGCGTCAA   | TTATTAAACA<br>AATAATTTGT  |
| 1501 | TAGCGTGGGA<br>ATCGCACCTT  | TCTCCGACAT<br>AGAGGCTGTA | CTCGGGTACG<br>GAGCCCATGC | TGTTCCGGAC<br>ACAAGGCTG   | ATGGGCTCTT<br>TACCCGAGAA  |
| 1551 | CTCCGGTAGC<br>GAGGCCATCG  | GGCGGAGCTT<br>CCGCCTCGAA | CCACATCCGA<br>GGTGTAGGCT | GCCCTGGTCC<br>CGGGACCAGG  | CATCCGTCCA<br>GTAGGCAGGT  |
| 1601 | GCGGCTCATG<br>CGCCGAGTAC  | GTCGCTCGGC<br>CAGCGAGCCG | AGCTCCTTGC<br>TCGAGGAACG | TCCTAACAGT<br>AGGATTGTCA  | GGAGGCCAGA<br>CCTCCGGTCT  |
| 1651 | CTTAGGCACA<br>GAATCCGTGT  | GCACAATGCC<br>CGTGTTACGG | CACCACCACC<br>GTGGTGGTGG | AGTGTGCCGC<br>TCACACGGCG  | ACAAGGCCGT<br>TGTTCCGGCA  |
| 1701 | GGCGGTAGGG<br>CCGCCATCCC  | TATGTGTCTG<br>ATACACAGAC | AAAATGAGCT<br>TTTTACTCGA | CGGAGATTGG<br>GCCTCTAACC  | GCTCGCACCT<br>CGAGCGTGGA  |
| 1751 | GGACGCAGAT<br>CCTGCGTCTA  | GGAAGACTTA<br>CCTTCTGAAT | AGGCAGCGGC<br>TCCGTCGCCG | AGAAGAAGAT<br>TCTTCTTCTA  | GCAGGCAGCT<br>CGTCCGTCGA  |
| 1801 | GAGTTGTTGT<br>CTCAACAACA  | ATTCTGATAA<br>TAAGACTATT | GAGTCAGAGG<br>CTCAGTCTCC | TAACTCCCGT<br>ATTGAGGGCA  | TGCGGTGCTG<br>ACGCCACGAC  |

FIG. 9-2



# pCMV-NS34A

```

1851 TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG
    AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC

1901 CGCCACCAGA CATAATAGCT GACAGACTAA CAGACTGTTC CTTTCCATGG
    GCGGTGGTCT GTATTATCGA CTGTCTGATT GTCTGACAAG GAAAGGTACC

+2                                     M A P
                                     EcoRI
                                     -----
1951 GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCACC ATGGCGCCCA
    CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTGG TACCGCGGGT

+2 I T A Y A Q Q T R G L L G C I I T
2001 TCACGGCGTA CGCCCAGCAG ACAAGGGGCC TCCTAGGGTG CATAATCACC
    AGTGCCGCAT GCGGGTCGTC TGTTCCTCCG AGGATCCAC GTATTAGTGG

+2 S L T G R D K N Q V E G E V Q I V
2051 AGCCTAACTG GCCGGGACAA AAACCAAGTG GAGGGTGAGG TCCAGATTGT
    TCGGATTGAC CGGCCCTGTT TTTGGTTCAC CTCCCACTCC AGGTCTAACA

+2 S T A A Q T F L A T C I N G V C
2101 GTCAACTGCT GCCCAAACCT TCCTGGCAAC GTGCATCAAT GGGGTGTGCT
    CAGTTGAEGA CGGGTTTGA AGGACCGTTG CACGTAGTTA CCCACACGA

+2 W T V Y H G A G T R T I A S P K G
2151 GGACTGTCTA CCACGGGGCC GGAACGAGGA CCATCGCGTC ACCCAAGGGT
    CCTGACAGAT GGTGCCCCGG CTTTGCTCCT GGTAGCGCAG TGGGTTCCCA

-2 P V I Q M Y T N V D Q D L V G W P
2201 CCTGTCTATCC AGATGTATAC CAATGTAGAC CAAGACCTTG TGGGCTGGCC
    GGACAGTAGG TCTACATATG GTTACATCTG GTTCTGGAAC ACCCGACCGG

+2 A S Q G T R S L T P C T C G S S
2251 CGCTTCGCAA GGTACCCGCT CATTGACACC CTGCACTTGC GGCTCCTCGG
    GCGAAGCGTT CCATGGGCGA GTAACGTGGG GACGTGAACG CCGAGGAGCC

+2 D L Y L V T R H A D V I P V R R R
2301 ACCTTTACCT GGTCACGAGG CACGCCGATG TCATTCCCGT GCGCCGGCGG
    TGGAAATGGA CCAGTGCTCC GTGCGGCTAC AGTAAGGGCA CGCGGCCGCC

+2 G D S R G S L L S P R P I S Y L K
2351 GGTGATAGCA GGGGCAGCCT GCTGTCGCCC CGGCCCATTT CCTACTTGAA
    CCACTATCGT CCCCCTCGGA CGACAGCGGG GCCGGGTAAA GGATGAACTT

+2 G S S G G P L L C P A G H A V G
2401 AGGCTCCTCG GGGGGTCCGC TGTTGTGCCC CGCGGGGCAC GCCGTGGGCA
    TCCGAGGAGC CCCCAGGCG ACAACACGGG GCGCCCCGTG CGGCACCCGT

+2 I F R A A V C T R G V A K A V D F
2451 TATTTAGGGC CGCGGTGTGC ACCCGTGGAG TGGCTAAGGC GGTGGACTTT
    ATAAATCCCG GCGCCACACG TGGGCACCTC ACCGATTCCG CCACCTGAAA

+2 I P V E N L E T T M R S P V F T D
2501 ATCCCTGTGG AGAACCTAGA GACAACCATG AGGTCCCCGG TGTTACGGA
    TAGGGACACC TCTTGGATCT CTGTTGGTAC TCCAGGGGCC ACAAGTGCTT
  
```

FIG. 9-3



pcMV-NS34A

+2 N S S P P V V P Q S F Q V A H L  
2551 TAACTCCTCT CCACCACTAG TGCCCCAGAG CTTCCAGGTG GCTCACCTCC  
ATTGAGGAGA GGTGGTCATC ACGGGGTCTC GAAGGTCCAC CGAGTGGAGG

---

+2 H A P T G S G K S T K V P A A Y A  
2601 ATGCTCCAC AGGCAGCGGC AAAAGCACCA AGGTCCCGGC TGCATATGCA  
TACGAGGGTG TCCGTCGCCG TTTTCGTGGT TCCAGGGCCG ACGTATACGT

---

+2 A Q G Y K V L V L N P S V A A T L  
2651 GCTCAGGGCT ATAAGGTGCT AGTACTCAAC CCCTCTGTTG CTGCAACACT  
CGAGTCCCGA TATTCCACGA TCATGAGTTG GGGAGACAAC GACGTTGTGA

---

+2 G F G A Y M S K A H G I D P N I  
2701 GGGCTTTGGT GCTTACATGT CCAAGGCTCA TGGGATCGAT CCTAACATCA  
CCCCAAACCA CGAATGTACA GGTTCGAGT ACCCTAGCTA GGATTGTAGT

---

+2 R T G V R T I T T G S P I T Y S T  
2751 GGACCGGGGT GAGAACAATT ACCACTGGCA GCCCCATCAC GTACTCCACC  
CCTGGCCCCA CTCTTGTTAA TGGTGACCGT CGGGGTAGTG CATGAGGTGG

---

+2 Y G K F L A D G G C S G G A Y D I  
2801 TACGGCAAGT TCCTTGCCGA CGGCGGGTGC TCGGGGGGCG CTTATGACAT  
ATGCCGTTCA AGGAACGGCT GCCGCCACG AGCCCCCGC GAATACTGTA

---

+2 I I C D E C H S T D A T S I L G  
2851 AATAATTTGT GACGAGTGCC ACTCCACGGA TGCCACATCC ATCTTGGGCA  
TTATTAAACA CTGCTCACGG TGAGGTGCCT ACGGTGTAGG TAGAACCCGT

---

+2 I G T V L D Q A E T A G A R L V V  
2901 TTGGCACTGT CTTGACCAA GCAGAGACTG CGGGGGCGAG ACTGGTTGTG  
AACCGTGACA GGAAGTGGT CGTCTCTGAC GCCCCGCTC TGACCAACAC

---

+2 L A T A T P P G S V T V P H P N I  
2951 CTCGCCACCG CCACCCCTCC GGGCTCCGTC ACTGTGCCCC ATCCCAACAT  
GAGCGGTGGC GGTGGGGAGG CCCGAGGCAG TGACACGGGG TAGGGTTGTA

---

+2 E E V A L S T T G E I P F Y G K  
3001 CGAGGAGGTT GCTCTGTCCA CCACCGGAGA GATCCCTTTT TACGGCAAGG  
GTCCTCCAA CGAGACAGGT GGTGGCCTCT CTAGGGAAAA ATGCCGTTCC

---

+2 A I P L E V I K G G R H L I F C H  
3051 CTATCCCCCT CGAAGTAATC AAGGGGGGGA GACATCTCAT CTTCTGTCTAT  
GATAGGGGGA GCTTCATTAG TTCCCCCCT CTGTAGAGTA GAAGACAGTA

---

+2 S K K K C D E L A A K L V A L G I  
3101 TCAAAGAAGA AGTGCGACGA ACTCGCCGCA AAGCTGGTCG CATTGGGCAT  
AGTTTCTTCT TCACGCTGCT TGAGCGGCGT TTCGACCAGC GTAACCCGTA

---

+2 N A V A Y Y R G L D V S V I P T  
3151 CAATGCCGTG GCCTACTACC GCGGTCTTGA CGTGTCCGTC ATCCCGACCA  
GTTACGGCAC CGGATGATGG CGCCAGAACT GCACAGGCAG TAGGGCTGGT

---

+2 S G D V V V V A T D A L M T G Y T  
3201 GCGGCGATGT TGTCGTCGTG GCAACCGATG CCCTCATGAC CGGCTATACC  
CGCCGCTACA ACAGCAGCAC CGTTGGCTAC GGGAGTACTG GCCGATATGG

FIG. 9-4



pCMV-NS34A

+2 G D F D S V I D C N T C V T Q T V  
3251 GGC GACTTCG ACTCGGTGAT AGACTGCAAT ACGTGTGTCA CCCAGACAGT  
CCGCTGAAGC TGAGCCACTA TCTGACGTTA TGCACACAGT GGGTCTGTCA

+2 D F S L D P T F T I E T I T L P  
3301 CGATTTCAGC CTTGACCCTA CTTTACCAT TGAGACAATC ACGTCCCCC  
GCTAAAGTCG GAACTGGGAT GGAAGTGGTA ACTCTGTTAG TGCAGAGGGG

+2 Q D A V S R T Q R R G R T G R G K  
3351 AAGATGCTGT CTCCCGCACT CAACGTCGGG GCAGGACTGG CAGGGGGAAG  
TTCTACGAGA GAGGGCGTGA GTTGCAGCCC CGTCCTGACC GTCCCCCTC

+2 P G I Y R F V A P G E R P S G M F  
3401 CCAGGCATCT ACAGATTTGT GGCACCGGGG GAGCGCCCCT CCGGCATGTT  
GGTCCGTAGA TGTCTAAACA CCGTGGCCCC CTCGCGGGGA GGCCGTACAA

+2 D S S V L C E C Y D A G C A W Y  
3451 CGACTCGTCC GTCCTCTGTG AGTGCTATGA CGCAGGCTGT GCTTGGTATG  
GCTGAGCAGG CAGGAGACAC TCACGATACT GCGTCCGACA CGAACCATAC

+2 E L T P A E T T V R L R A Y M N T  
3501 AGCTCACGCC CGCCGAGACT ACAGTTAGGC TACGAGCGTA CATGAACACC  
TCGAGTGCGG GCGGCTCTGA TGTCAATCCG ATGCTCGCAT GTACTTGTGG

+2 P G L P V C Q D H L E F W E G V F  
3551 CCGGGGCTTC CCGTGTGCCA GGACCATCTT GAATTTTGGG AGGGCGTCTT  
GGCCCCGAAG GGCACACGGT CCTGGTAGAA CTTAAAACCC TCCCGCAGAA

+2 T G L T H I D A H F L S Q T K Q  
StuI  
~~~~~

3601 TACAGGCCTC ACTCATATAG ATGCCCACTT TCTATCCCAG ACAAAGCAGA  
ATGTCCGGAG TGAGTATATC TACGGGTGAA AGATAGGGTC TGTTTCGTCT

+2 S G E N L P Y L V A Y Q A T V C A  
3651 GTGGGGAGAA CCTTCCTTAC CTGGTAGCGT ACCAAGCCAC CGTGTGCGCT  
CACCCCTCTT GGAAGGAATG GACCATCGCA TGGTTCGGTG GCACACGCGA

+2 R A Q A P P P S W D Q M W K C L I  
3701 AGGGCTCAAG CCCCTCCCC ATCGTGGGAC CAGATGTGGA AGTGTGTGAT  
TCCCGAGTTC GGGGAGGGG TAGCACCTG GTCTACACCT TCACAACTA

+2 R L K P T L H G P T P L L Y R L  
3751 TCGCCTCAAG CCCACCCTCC ATGGGCCAAC ACCCCTGCTA TACAGACTGG  
AGCGGAGTTC GGGTGGGAGG TACCCGGTTG TGGGGACGAT ATGTCTGACC

+2 G A V Q N E I T L T H P V T K Y I  
3801 GCGCTGTTCA GAATGAAATC ACCCTGACGC ACCCAGTCAC CAAATACATC  
CGCGACAAGT CTTACTTTAG TGGGACTGCG TGGGTCAGTG GTTTATGTAG

+2 M T C M S A D L E V V T S T W V L  
3851 ATGACATGCA TGTCGGCCGA CCTGGAGGTC GTCACGAGCA CCTGGGTGCT  
TACTGTACGT ACAGCCGGCT GGACCTCCAG CAGTGCTCGT GGACCCACGA

+2 V G G V L A A L A A Y C L S T G  
3901 CGTTGGCGGC GTCCTGGCTG CTTTGGCCGC GTATTGCTG TCAACAGGCT  
GCAACCGCCG CAGGACCGAC GAAACCGGCG CATAACGGAC AGTTGTCCGA

FIG. 9-5



# pCMV-NS34A

	+2	C	V	V	I	V	G	R	V	V	L	S	G	K	P	A	I	I									
3951		GC	TG	GC	AT	AG	TG	GG	GC	AG	GT	CG	TCT	TGT	CC	GG	GA	AG	CC								
		CG	CA	CC	AG	TA	CA	CC	CG	TCC	CA	GC	AGA	AACA	GG	CC	CT	TC	CG								
	+2	P	D	R	E	V	L	Y	R	E	F	D	E	M	E	E	C										
4001		CCT	GAC	AGG	G	AAG	TCT	CTA	CCG	AG	AGT	TTC	GAT	GAG	ATG	G	AAG	AGT	GC	TA							
		GG	ACT	GT	CCC	TT	CAG	GAG	AT	GG	CT	CT	CA	AG	CT	ACT	CT	ACC	TT	CT	CAC	GC	AT				
		Bam	HI		Mlu	I																					
		~~~~~			~~~~~																						
4051		GG	AT	CC	ACT	TA	GAG	CT	CG	CT	GAT	C	AG	CC	TC	GACT	GT	GC	CT	TCT	TA						
		CCT	AG	GT	GAT	GCG	CAAT	CTC	GAG	CG	ACT	AG	TC	GG	AG	CT	GA	CAC	GG	AAG	AT						
4101		GT	TGCC	AG	CC	AT	CT	GT	TG	TT	GT	TG	CC	CC	CT	CCC	CC	GT	GC	CT	TG						
		CA	AC	GG	TC	GG	TAG	ACA	ACA	AA	AC	GG	GG	GAG	GG		GG	CAC	GG	AAG	GA	ACT	GG	GAC			
4151		GA	AGGT	GC	CA	CT	CC	CA	CT	GT	CCT	TT	CT	CT	AA	TAAA	AT	GAG	G	AA	AT	TG	CA	TC			
		CT	TC	CA	CG	GT	GAG	GG	TG	ACA	GG	AA	AG	GAT	T	AT	TT	TACT	CC	TT	TA	AC	GT	AG			
4201		GC	ATT	GT	CT	GT	AG	TAG	GT	GT	C	ATT	CT	ATT	CT	GG	GG	GG	TG	GG	GT	GG	GG	CAG	G		
		CG	TA	AC	AG	AC	TC	AT	CC	AC	AG	TA	AG	ATA	AGA	CC	CC	CC	CACC	CAC	CC	CG	TC	CC			
4251		AC	AG	CA	AG	GG	GG	AG	GAT	TG	G	GA	AG	CA	ATA	GC	AG	GC	AT	GC	TG	GG	GG	GAG	CT		
		TG	TC	GT	TCCC	CCT	CCT	TA	ACC		CT	TCT	GT	TAT	CG	TCC	GT	ACG	AC	CC	CT	CG	AG				
4301		TT	CC	GC	TT	CC	TC	GC	TACT	GC	ACT	CG	CT	GC	G	CT	CG	GT	CG	TT	CG	GT	GC	GG	C		
		AAG	GC	GA	AG	G	AG	CG	AG	TG	AC	TG	AG	CG	AC	GC	GAG	CC	AG	CAA		GC	CC	GAC	CC		
4351		GAG	CG	GT	AT	C	AG	CT	CA	CT	CA	AAG	GC	GG	TAA	TAC	GG	TT	AT	C	CAC	AGA	AT	CA	TA		
		CT	CG	CC	AT	AG	TC	GAG	TG	AG	T	TT	CC	GC	CATT	AT	GCC	AAT	AG		GT	GT	CT	TAG			
4401		GG	GG	ATA	AC	G	CAG	GAA	AG	AA	CAT	GT	GAG	CA	AA	AG	GC	CAG	C	AAA	AG	GC	CA	G			
		CCC	CT	ATT	TG	C	GT	C	TTT	CT	T	GT	AC	ACT	CG	TT	CC	GG	TC	TT	TT	CC	GG	TC			
4451		GA	ACC	GT	AAA	AAG	GC	CG	CG	T	TG	CT	GG	CG	T	TT	T	CC	AT	AG	CT	CC	GC	CCCC			
		CT	TG	GC	ATT	T	TT	CC	GG	CG	CA	AC	G	ACC	GC	AA	AA	AG	GT	AT	CC	GAG	GC	GG	GG		
4501		CT	GAC	GAG	CA	TC	AC	AAAA	AT		CG	AC	GC	TCA	A	GT	CAG	AG	GT	G	GCG	AA	AC	CCG			
		GAC	TG	CT	CG	T	AG	TG	TTTT	TA	GCT	GC	GAG	TT	CAG	T	CT	CC	AC	CG	CT	TT	GG	GC			
4551		AC	AG	GA	CT	AT	AA	AG	ATA	CCA	GG	CG	TTT	CCC	CCT	GGA	AG	CT	CC	CT	CG	TG	CG				
		TG	T	CT	GATA	TTT	CT	AT	GG	T	CC	G	CAA	AG	GG	GG	AC	CT	TC	GA	GG	GAG	CAC	GC			
4601		CT	CT	C	TG	T	CC	G	ACC	CT	GC	CG	CT	TAC	CG	G	AT	AC	CT	GT	CC	GC	CT	TT	CT	CC	
		GAG	AG	G	ACAA	GG	CT	GGG	ACG		GCG	AAT	GG	CC	TAT	GG	AC	AG		CG	GAA	AG	AG				
4651		CT	T	CG	GA	AG	CG	T	GG	CG	CT	TCT	CA	AT	GC	T	CAC	GC	T	GT	AG	GT	AT	CT	CAG	T	
		GA	AG	CC	CT	T	GC	ACC	GC	GA	AG	AG	TT	AC	GA	GT	GC	GAC	AT	C	CAT	AG	AG	TCA			
4701		TC	GG	T	G	TAG	TC	GG	T	TC	GC	T	C	CA	AG	CT	GG	G	TG	T	G	TG	CAC	G			
		AG	CC	AC	AT	CC	AG	CA	AG	CG	AG	GT	TC	G	AC	CC	G	AC	AC	AC	GT	GC	TT	GG	GG	GG	GCA
4751		TC	AG	CC	CG	AC	CG	CT	G	CG	CT	TAT	CC	GG	TAA	CT	AT	CG	T	CT	GAG	T	CCA	ACC			
		AG	T	CG	GG	CT	G	CG	AC	CG	GA	AT	AG	GG	CC	ATT	GAT	AG	CAG	AA	CT	CAG	GT	TG			
4801		CG	G	TAA	GACA	CG	ACT	TAT	CG		CC	ACT	G	G	CAG	CAG	CC	ACT	G	G	TA	AC	AG	GAT	T		
		GCC	ATT	CT	GT	GCT	GA	AT	AG	C	GG	T	G	AC	CG	T	GTC	GG	T	G	AC	ATT	GT	C	CTAA		

FIG. 9-6





pCMV-NS34A

4851	AGCAGAGCGA TCGTCTCGCT	GGTATGTAGG CCATACATCC	CGGTGCTACA GCCACGATGT	GAGTTCTTGA CTCAAGAACT	AGTGGTGGCC TCACCACCGG
4901	TAACTACGGC ATTGATGCCG	TACACTAGAA ATGTGATCTT	GGACAGTATT CCTGTCATAA	TGGTATCTGC ACCATAGACG	GCTCTGCTGA CGAGACGACT
4951	AGCCAGTTAC TCGGTCAATG	CTTCGGAAAA GAAGCCTTTT	AGAGTTGGTA TCTCAACCAT	GCTCTTGATC CGAGAACTAG	CGGCAAACAA GCCGTTTGTT
5001	ACCACCGCTG TGGTGGCGAC	GTAGCGGTGG CATCGCCACC	TTTTTTTGTT AAAAAAACAA	TGCAAGCAGC ACGTTTCGTCG	AGATTACGCG TCTAATGCGC
5051	CAGAAAAAAA GTCTTTTTTT	GGATCTCAAG CCTAGAGTTC	AAGATCCTTT TTCTAGGAAA	GATCTTTTCT CTAGAAAAGA	ACGGGGTCTG TGCCCCAGAC
5101	ACGCTCAGTG TGCGAGTCAC	GAACGAAAAC CTTGCTTTTG	TCACGTAAAG AGTGCAATTC	GGATTTTGGT CCTAAAACCA	CATGAGATTA GTACTCTAAT
5151	TCAAAAAGGA AGTTTTTCCT	TCTTCACCTA AGAAGTGGAT	GATCCTTTTA CTAGGAAAAAT	AATTAAAAAT TTAATTTTTA	GAAGTTTAA CTTCAAAAT
5201	ATCAATCTAA TAGTTAGATT	AGTATATATG TCATATATAC	AGTAAACTTG TCATTTGAAC	GTCTGACAGT CAGACTGTCA	TACCAATGCT ATGGTTACGA
5251	TAATCAGTGA ATTAGTCACT	GGCACCTATC CCGTGGATAG	TCAGCGATCT AGTCGCTAGA	GTCTATTTTCG CAGATAAAGC	TTCATCCATA AAGTAGGTAT
5301	GTTGCCTGAC CAACGGACTG	TCCCCGTCGT AGGGGCAGCA	GTAGATAACT CATCTATTGA	ACGATACGGG TGCTATGCCC	AGGGCTTACC TCCCGAATGG
5351	ATCTGGCCCC TAGACCGGGG	AGTGCTGCAA TCACGACGTT	TGATACCGCG ACTATGGCGC	AGACCCACGC TCTGGGTGCG	TCACCGGCTC AGTGGCCGAG
5401	CAGATTTATC GTCTAAATAG	AGCAATAAAC TCGTTATTTG	CAGCCAGCCG GTCGGTCGGC	GAAGGGCCGA CTTCCCGGCT	GCGCAGAAGT CGCGTCTTCA
5451	GGTCCTGCAA CCAGGACGTT	CTTTATCCGC GAAATAGGCG	CTCCATCCAG GAGGTAGGTC	TCTATTAATT AGATAATTAA	GTTGCCGGGA CAACGGCCCT
5501	AGCTAGAGTA TCGATCTCAT	AGTAGTTCGC TCATCAAGCG	CAGTTAATAG GTCAATTATC	TTTGCGCAAC AAACGCGTTG	GTTGTTGCCA CAACAACGGT
5551	TTGCTACAGG AACGATGTCC	CATCGTGGTG GTAGCACCAC	TCACGCTCGT AGTGCAGAGCA	CGTTTGGTAT GCAAACCATA	GGCTTCATTC CCGAAGTAAG
5601	AGCTCCGGTT TCGAGGCCAA	CCCAACGATC GGGTTGCTAG	AAGGCGAGTT TTCCGCTCAA	ACATGATCCC TGTAAGTAGG	CCATGTTGTG GGTACAACAC
5651	CAAAAAAGCG GTTTTTTCGC	GTTAGCTCCT CAATCGAGGA	TCGGTCCTCC AGCCAGGAGG	GATCGTTGTC CTAGCAACAG	AGAAGTAAGT TCTTCATTCA
5701	TGGCCGCGAGT ACCGGCGTCA	GTTATCACTC CAATAGTGAG	ATGGTTATGG TACCAATACC	CAGCACTGCA GTCGTGACGT	TAATTCTCTT ATTAAGAGAA
5751	ACTGTCATGC TGACAGTACG	CATCCGTAAG GTAGGCATTC	ATGCTTTTCT TACGAAAAGA	GTGACTGGTG CACTGACCAC	AGTACTCAAC TCATGAGTTG

FIG. 9-7



pCMV-NS34A

5801	CAAGTCATT	CGAGTAAG	ACTCTTAT	CAATACG	CGCCG	ACCGAGTT	GCTTGCCC	GGG
5851	CGTCAATAC	GGATAATAC	GCGCCACAT	GCAGAACTTT	AAAAGTGCTC	GCAGTTATG	CCTATTATG	CGCGGTGTAT
5901	ATCATTGGAA	AACGTTCTTC	GGGGCGAAAA	CTCTCAAGGA	TCTTACCGCT	TAGTAACCTT	TTGCAAGAAG	CCCCGCTTTT
5951	GTTGAGATCC	AGTTCGATGT	AACCCACTCG	TGCACCCAAC	TGATCTTCAG	CAACTCTAGG	TCAAGCTACA	TTGGGTGAGC
6001	CATCTTTTAC	TTTCAACCAGC	GTTTCTGGGT	GAGCAAAAAC	AGGAAGGCAA	GTAGAAAATG	AAAGTGGTCG	CAAAGACCCA
6051	AATGCCGCAA	AAAAGGGAAT	AAGGGCGACA	CGGAAATGTT	GAATACTCAT	TTACGGCGTT	TTTTCCCTTA	TTCCCGCTGT
6101	ACTCTTCCTT	TTTCAATATT	ATTGAAGCAT	TTATCAGGGT	TATTGTCTCA	TGAGAAGGAA	AAAGTTATAA	TAACTTCGTA
6151	TGAGCGGATA	CATATTTGAA	TGTATTTAGA	AAAATAAACA	AATAGGGGTT	ACTCGCCTAT	GTATAAACTT	ACATAAATCT
6201	CCGCGCACAT	TTCCCCGAAA	AGTGCCACCT	GACGTCTAAG	AAACCATTAT	GGCGCGTGTA	AAGGGGCTTT	TCACGGTGGA
6251	TATCATGACA	TTAACCTATA	AAAATAGGCG	TATCACGAGG	CCCTTTCGTC	ATAGTACTGT	AATTGGATAT	TTTTATCCGC

FIG. 9-8

Diagram 1

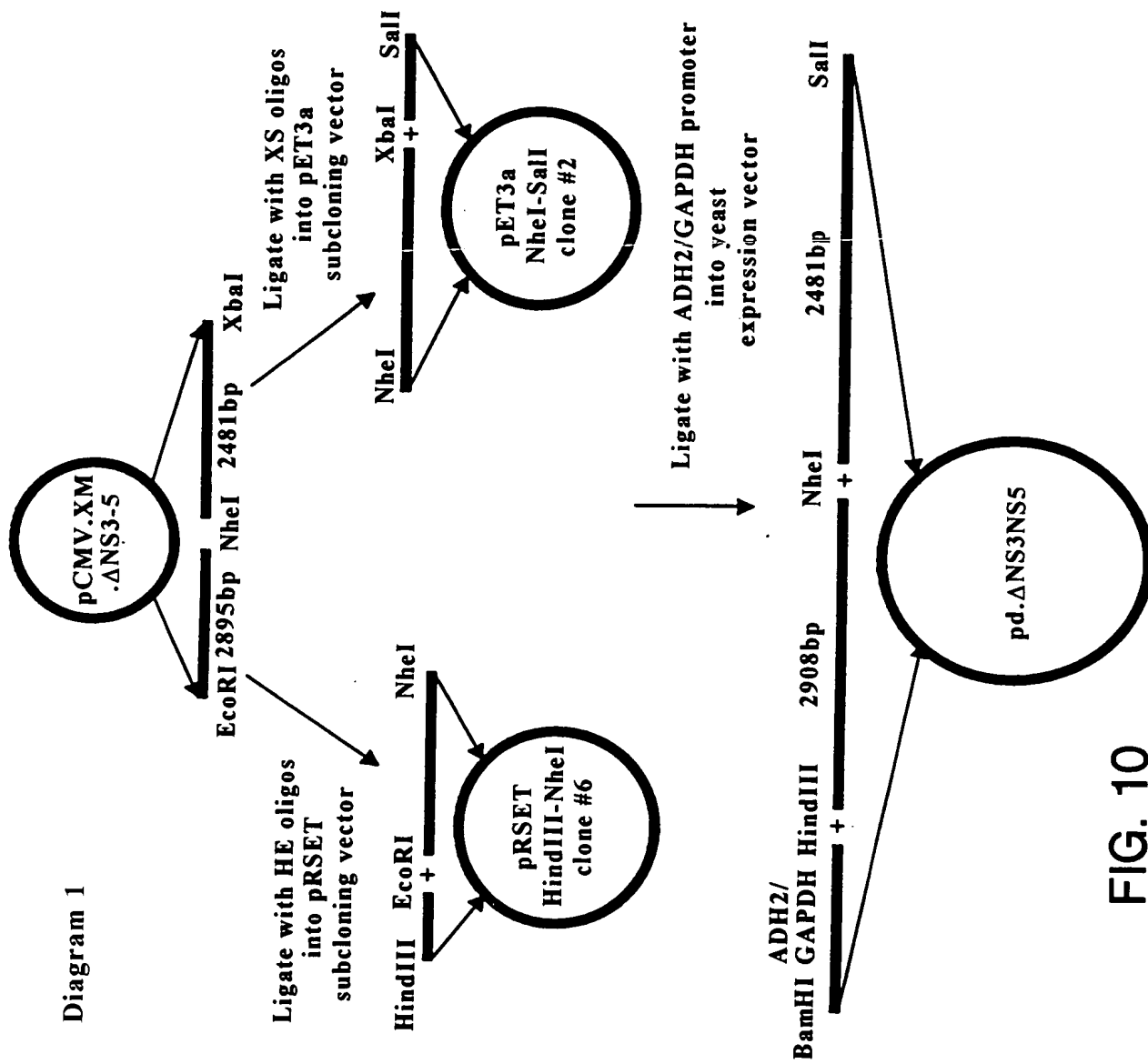


FIG. 10



MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuVal  
2 AGCTTACAAAACAAATTCACCATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTA  
TCGAATGTTTTGTTTAAGTGGTACCGACGTATACGTCGAGTCCCGATATTCACGATCAT  
^ ^ ^  
1 HIND3, 21 NCOI, 30 NDEI, 58 SCAI,  
  
LeuAsnProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGly  
62 CTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGG  
GAGTTGGGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCC  
  
IleAspProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyr  
122 ATCGATCCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATCACGTAC  
TAGCTAGGATTGTAGTCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATG  
^  
122 CLAI,  
  
SerThrTyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIle  
182 TCCACCTACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATA  
AGGTGGATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGGAATACTGTATTAT  
  
IleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeu  
242 ATTTGTGACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTT  
TAAACACTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAA  
  
AspGlnAlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGly  
302 GACCAAGCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGC  
CTGGTTCGTCTCTGACGCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCC  
^  
309 ALWN1,  
  
SerValThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIle  
362 TCCGTCACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATC  
AGGCAGTGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAG  
  
ProPheTyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePhe  
422 CCTTTTTACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGGAGACATCTCATCTTC  
GGAAAATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAG  
  
CysHisSerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsn  
482 TGTCATTCAAAGAAGAAGTGCACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAAT  
ACAGTAAGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTA  
  
AlaValAlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValVal  
542 GCCGTGGCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCGCGCGATGTTGTC  
CGGCACCGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCGCTACAACAG  
^ ^  
556 SAC2, 566 DRD1,  
  
ValValAlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAsp  
602 GTCGTGGCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGAC  
CAGCACCGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTG  
^  
621 BSPH1,  
  
CysAsnThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGlu

FIG. 11-1



662 TGCAATACGTGTGTCACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAG  
ACGTTATGCACACAGTGGGTCTGTCAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTC

722 ThrIleThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArg  
ACAATCACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGG  
TGTTAGTGCAGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCC

782 GlyLysProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAsp  
GGGAAGCCAGGCATCTACAGATTTGTGGCACCAGGGGGAGCGCCCCCTCCGGCATGTTTCGAC  
CCCTTCGGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTG  
822 BGLI, 839 DRD1,

842 SerSerValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAla  
TCGTCCGTCCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCC  
AGCAGGCAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGG  
887 SACI,

902 GluThrThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAsp  
GAGACTACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGAC  
CTCTGATGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTCTG  
937 SMAI XMAI,

962 HisLeuGluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeu  
CATCTTGAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTA  
GTAGAACTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGAT  
991 STUI,

1022 SerGlnThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrVal  
TCCCAGACAAAGCAGAGTGGGGAGAACCTTCCCTTACCTGGTAGCGTACCAAGCCACCGTG  
AGGGTCTGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCAC  
1075 DRA3,

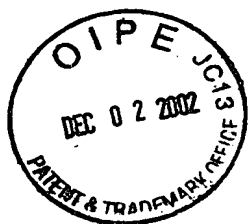
1082 CysAlaArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArg  
TGCGCTAGGGCTCAAGCCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTCCG  
ACGCGATCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTCACAACTAAGCG

1142 LeuLysProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsn  
CTCAAGCCCACCCTCCATGGGGCAACACCCTGCTATACAGACTGGGCGCTGTTTCAGAAT  
GAGTTCGGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTA  
1156 NCOI,

1202 GluIleThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeu  
GAAATCACCTGACGCACCCAGTCACCAATACATCATGACATGCATGTCGGCCGACCTG  
CTTTAGTGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGAC  
1236 BSPH1, 1240 DRD1, 1243 AVA3, 1251 EAG1 XMA3, 1256 DRD1,

1262 GluValValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyr  
GAGGTGCTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTAT  
CTCCAGCAGTGCTCGTGACCCACGAGCAACCGCCGAGGACCGACGAACCGGCGCATA

FIG. 11-2



1322 CysLeuSerThrGlyCysValValIleValGlyArgValValL uSerGlyLysPr Ala  
TGCCTGTCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCA  
ACGGACAGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGT  
1375 NAEI,

1382 IleIleProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGln  
ATCATACCTGACAGGGAAGTCCTCTACCGAGAGTTTCGATGAGATGGAAGAGTGCTCTCAG  
TAGTATGGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTC  
1391 DRD1,

1442 HisLeuProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeu  
CACTTACCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTC  
GTGAATGGCATGTAGCTCGTTCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAG

1502 GlyLeuLeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsn  
GGCCTCCTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAAC  
CCGGAGGACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTG  
1508 PSTI, 1513 TTH3I,

1562 TrpGlnLysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGln  
TGGCAAAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAA  
ACCGTTTTTGAGCTCTGGAAGACCCGCTTCGTATACACCTGAAGTAGTCACCCTATGTT  
1571 XHOI, 1592 NDEI,

1622 TyrLeuAlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPhe  
TACTTGGCGGGCTTGTCACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTT  
ATGAACCGCCCGAACAGTTGCGACGGACCATTTGGGGCGGTAACGAAGTAACACCGAAAA  
1649 BSTE2,

1682 ThrAlaAlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGly  
ACAGCTGCTGTCACCAGCCCACTAACCCTAGCCAAACCCTCCTCTTCAACATATTGGGG  
TGTCGACGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCC  
1683 ALWN1 PVU2,

1742 GlyTrpValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGly  
GGGTGGGTGGCTGCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGC  
CCCACCCACCGACGGGTGAGCGGGCGGGGCCACGGCGATGACGGAAACACCCGCGACCG  
1800 ESP1,

1802 LeuAlaGlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAla  
TTAGCTGGCGCCGCCATCGGCAGTGTGGACTGGGGAAGGTCCTCATAGACATCCTTGCA  
AATCGACCGCGGGCGGTAGCCGTACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGT  
1808 KAS1 NARI,

1862 GlyTyrGlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluVal  
GGGTATGGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTC  
CCCATACCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAG

FIG. 11-3



1884 SACI, 1905 BSPH1,

1922 ProSerThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaL uVal  
CCCTCCACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCCGAGCCCTCGTA  
GGGAGGTGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCAT  
^

1934 TTH3I,

1982 ValGlyValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaVal  
GTCGGCGTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTG  
CAGCCGCACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCCGCTCCCCCGTCAC  
^ ^

2010 NAEI, 2023 SMAI XMAI,

2042 GlnTrpMetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHis  
CAGTGGATGAACCGGCTGATAGCCTTCGCCTCCCGGGGGAACCATGTTTCCCCACGCAC  
GTCACCTACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTG  
^ ^

2073 SMAI XMAI, 2099 DRA3,

2102 TyrValProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrVal  
TACGTGCCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTA  
ATGCACGGCCTCTCGCTACGTGACGGGCGCAGTGACGGTATGAGTCGTGCGAGTGACAT  
^

2121 PVU2,

2162 ThrGlnLeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSer  
ACCCAGCTCCTGAGGCGACTGCACCACTGGATAAGCTCGGAGTGTAACCTCCATGCTCC  
TGGGTGAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGG  
^ ^

2165 ALWN1, 2170 MST2,

2222 GlySerTrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThr  
GGTTCCTGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACC  
CCAAGGACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGTGAAATTCTGG  
^

2226 ECON1,

2282 TrpLeuLysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArg  
TGGCTAAAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCCTTGTGTCTGCCAGCGC  
ACCGATTTTCGATTGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTCGCG  
^ ^ ^

2291 ESP1, 2306 PVU2, 2316 BAMHI,

2342 GlyTyrLysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAla  
GGGTATAAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCT  
CCCATATTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGA

2402 GluIleThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArg  
GAGATCACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGG  
CTCTAGTGACCTGTACAGTTTTTGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCC  
^ ^ ^

2431 BSAB1, 2447 AVR2, 2454 SSE83871, 2455 PSTI,

2462 AsnMetTrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyPr CysThrProLeu  
AACATGTGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGGCCCCTGTACCCCCCTT  
TTGTACACCTCACCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGGAA

FIG. 11-4



2486 ASE1, 2503 APAI,

2522 ProAlaProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIle  
CCTGCGCCGAACACTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATA  
GGACGCGGCTTGATGTGCAAGCGGATACCTCCCACAGACGTCTCCTTATGCACCTCTAT  
^

2559 PSTI,

2582 ArgGlnValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysPro  
AGGCAGGTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCCG  
TCCGTCCACCCCTGAAGGTGATGACTGCCCATCTGATGACTGTTAGAATTTACGGGC  
^

2600 DRA3,

2642 CysGlnValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPhe  
TGCCAGGTCCCATCGCCCGAATTTTTTCACAGAATTGGACGGGGTGCCTACATAGGTTT  
ACGGTCCAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAA

2702 AlaProProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGlu  
GCGCCCCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTAGAGTAGGACTCCACGAA  
CGCGGGGGACGTTGCGGAACGACGCCCTCCTCCATAGTAAGTCTCATCTGAGGTGCTT

2762 TyrProValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSer  
TACCCGGTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCC  
ATGGGCCATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACCTGCAGG  
^

2763 HGIE2, 2815 AAT2,

2822 MetLeuThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGly  
ATGCTCACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGGCGAAGGTTGGCGAGGGGA  
TACGAGTGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCGCTTCCAACCGCTCCCCT  
^

2856 EAG1 XMA3,

2882 SerProProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAla  
TCACCCCCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCA  
AGTGGGGGGAGACACCGGTCGAGGAGCCGATCGGTGATAGGCGAGGTAGAGAGTTCCGT  
^

2895 BALI, 2909 NHEI,

2942 ThrCysThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrp  
ACTTGCACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGG  
TGAACGTGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACC  
^ ^

2972 ESP1, 2975 SACI,

3002 ArgGlnGluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeu  
AGGCAGGAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTG  
TCCGTCTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTGTTTCACCACTAAGAC

3062 AspSerPheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGlu  
GACTCCTTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAA  
CTGAGGAAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTT  
^

3102 BGL2,

FIG. 11-5





3122 IleLeuArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyr  
ATCCTGCGGAAGTCTCGGAGATTGCGCCAGGCCCTGCCCGTTTGGGCGCGGCCGGACTAT  
TAGGACGCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATA  
3149 ALWN1, 3170 EAG1 XMA3,  
3182 AsnProProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGly  
AACCCCCCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGC  
TTGGGGGGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCG  
3223 HGIE2, 3235 NCOI,  
3242 CysProLeuProProProLysSerProProValProProProArgLysLysArgThrVal  
TGCCCGCTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTG  
ACGGGCGAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCAC  
3302 ValLeuThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGly  
GTCCTCACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGC  
CAGGAGTGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCG  
3338 SACI, 3352 HIND3,  
3362 SerSerSerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaPro  
AGCTCCTCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCT  
TCGAGGAGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGA  
3422 SerGlyCysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGly  
TCTGGCTGCCCCCGACTCCGACGCTGAGTCTATTCTCCATGCCCCCTGGAGGGG  
AGACCGACGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCC  
3443 EAM11051,  
3482 GluProGlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsn  
GAGCCTGGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAAC  
CTCGGACCCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGGTTG  
3490 BAMHI, 3491 BSAB1, 3493 BSPE1,  
3542 AlaGluAspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrPro  
GCGGAGGATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCG  
CGCCTCTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGC  
3595 DRA3,  
3602 CysAlaAlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHis  
TGCGCCGCGGAAGAACAGAAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCAC  
ACGCGGCGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTCTGTTAGCAACGATGCAGTG  
3606 SAC2, 3617 ALWN1, 3661 PFLM1,  
3662 HisAsnLeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThr  
CACAAATTTGGTGTATTCCACCACCTCACGAGTGCTTGCCAAAGGCAGAAAGTACACA  
GTGTTAAACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGTCTTCTTTCAGTGT  
3687 DRA3,  
PheAspArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAla

FIG. 11-6



3722 TTTGACAGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCA  
AAACTGTCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTTCTGT

3782 AlaAlaSerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrPro  
GCGGCGTCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACGCCTGACGCCC  
CGCCGCAGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTGCGACTGCGGG

3822 HIND3,

3842 ProHisSerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArg  
CCACACTCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGA  
GGTGTGAGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCT

3881 AAT2, 3896 BGLI,

3902 LysAlaValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrPro  
AAGGCCGTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCA  
TTCCGGCATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGT

3962 IleAspThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGly  
ATAGACACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGT  
TATCTGTGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCA

4022 ArgLysProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMet  
CGTAAGCCAGCTCGTCTCATCGTGTTCCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATG  
GCATTCCGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTAC

4082 AlaLeuTyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPhe  
GCTTTGTACGACGTGGTTACAAAGCTCCCCCTGGCCGTGATGGGAAGCTCCTACGGATTCT  
CGAAACATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCTTCGAGGATGCCTAAG

4142 GlnTyrSerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThr  
CAATACTCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACC  
GTTATGAGTGGTCTGTGCGCCCACTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTTGG

4166 ECORI,

4202 ProMetGlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIle  
CCAATGGGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATC  
GGTTACCCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCAGTGAATCTCGCTGTAG

4235 DRD1, 4242 ALWN1,

4262 ArgThrGluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIle  
CGTACGGAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCAAGCCCGGTGGCCATC  
GCATGCCTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCCGGCGCACCGGTAG

4307 BGLI, 4314 BALI,

4322 LysSerLeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsn  
AAGTCCCTCACCAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAAC  
TTCAGGGAGTGGCTCTCCGAAATACAACCCCGGGAGAATGGTTAAGTTCCCCCTCTTG

4351 APAI,

4382 CysGlyTyrArgArgCysArgAlaS rGlyValLeuThrThrSerCysGlyAsnThrLeu  
TGCGGCTATCGCAGGTGCCGCGGAGCGGCGTACTGACAACTAGCTGTGGTAACACCTC

FIG. 11-7



ACGCCGATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAG

4442 ThrCysTyrI1 LysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMet  
ACTTGCTACATCAAGGCCCCGGGCAGCCTGTGAGCCGCAGGGCTCCAGGACTGCACCATG  
TGAACGATGTAGTTCCGGGCCCCGTCGGACAGCTCGGCGTCCCGAGGTCTGACGTGGTAC  
^

4458 SMAI XMAI,

4502 LeuValCysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAla  
CTCGTGTGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGTCCAGGAGGACGCG  
GAGCACACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCTCTCTGCGC  
^ ^

4514 DRD1, 4517 TTH3I,

4562 AlaSerLeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspPro  
GCGAGCCTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCTGGGGACCCC  
CGCTCGGACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGG

4622 ProGlnProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAla  
CCACAACCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCC  
GGTGTGGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGG  
^

4643 SACI,

4682 HisAspGlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAla  
CACGACGGCGCTGGAAAGAGGGTCTACTACCTCACCGGTGACCCTACAACCCCCCTCGCG  
GTGCTGCCGCGACCTTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGC  
^

4737 NRUI,

4742 ArgAlaAlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIle  
AGAGCTGCGTGGGAGACAGCAAGACACACTCCAGTCAATTCTGGCTAGGCAACATAATC  
TCTCGACGCACCCTCTGTCTGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAG

4802 MetPheAlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeu  
ATGTTTGCCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTT  
TACAAACGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAA  
^^

4812 PFLM1, 4813 DRA3,

4862 IleAlaArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSer  
ATAGCCAGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCC  
TATCGGTCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGG  
^

4899 BGL2,

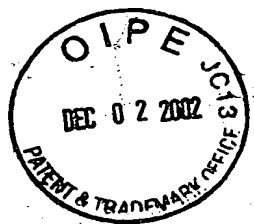
4922 IleGluProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSer  
ATAGAACCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCA  
TATCTTGGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGT  
^

4960 NCOI,

4982 LeuHisSerTyrSerPr GlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGly  
CTCCACAGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTTGGG  
GAGGTGTCAATGAGAGGTCCACTTTAGTTATCCCACCGCGTACGGAGTCTTTTGAACCC  
^

5021 SPHI, 5041 KPNI,

FIG. 11-8



5042 ValProProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAla  
GTACCGCCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCC  
CATGGCGGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGG  
5070 APAI, 5097 BALI,  
ArgGlyGlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLys  
5102 AGAGGAGGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAG  
TCTCCTCCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTTC  
5119 NDEI,  
LeuLysLeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAla  
5162 CTCAAACTCACTCCAATAGCGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCT  
GAGTTTGAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGA  
5180 NOTI, 5181 EAG1 XMA3, 5188 BALI, 5192 PVU2,  
GlyTyrSerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrp  
5222 GGCTACAGCGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCCGCTGGATCTGG  
CCGATGTGCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCGGGGCGACCTAGACC  
5246 DRA3,  
PheCysLeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgOP  
5282 TTTTGCCTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACCGATGAAGG  
AAAACGGATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTGGCTACTTCC  
5301 PSTI, 5331 HGIE2,  
5342 TTGGGGTAAACACTCCGGCCTAAAAAATACTAGAACCCGAGTCGAC  
AACCCCATTTGTGAGGCCGATTTTTTTTTTTTTTTAGATCTTGGGCTCAGCTG  
5378 XBAI, 5390 SALI,

FIG. 11-9

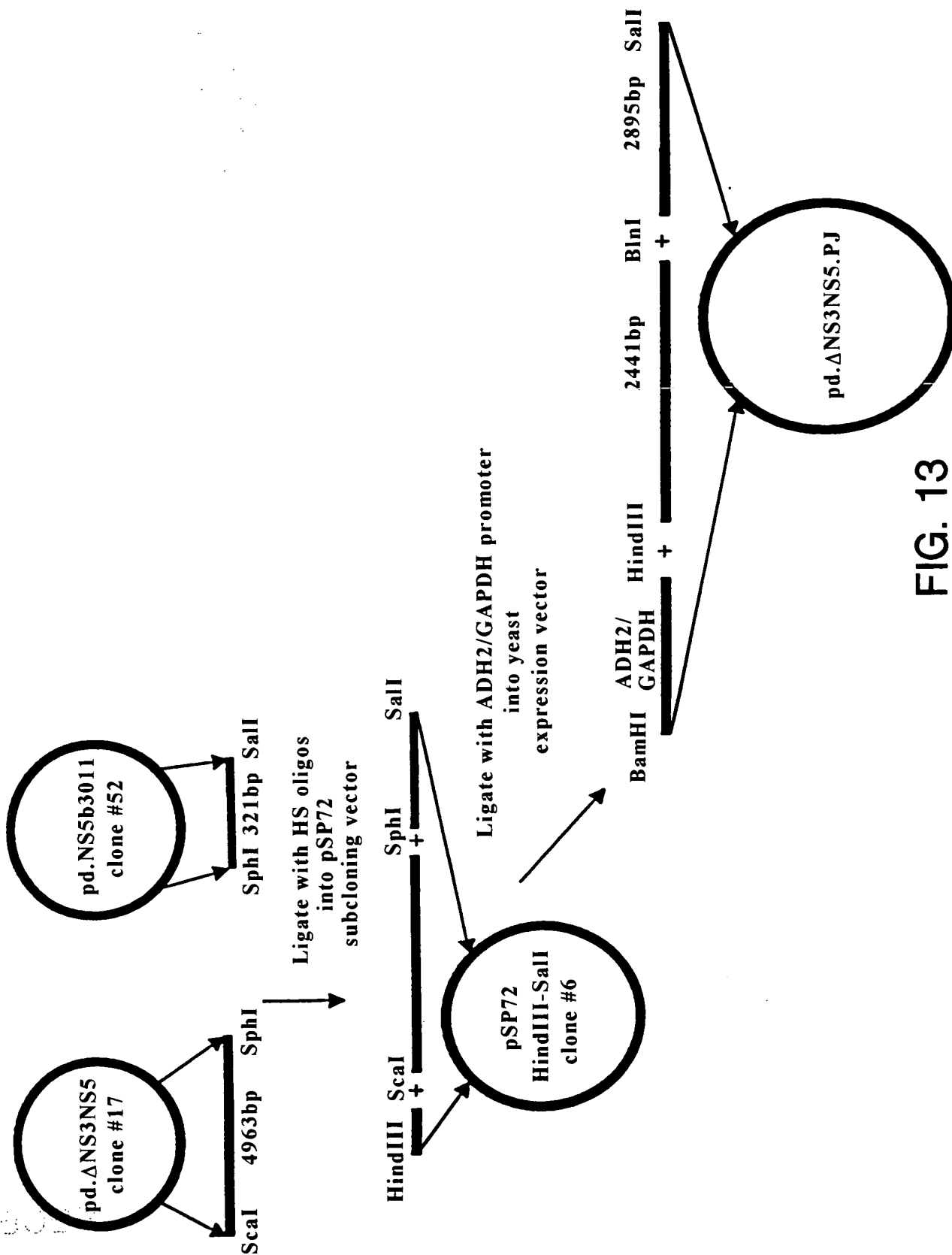
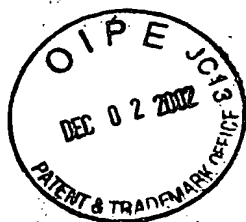


FIG. 13



MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTTGTTTTACCGACGTATACGTGAGTCCCGATATTCCACGATCATGAGTTG  
^ ^ ^  
1 HIND3, 24 NDEI, 52 SCAI,  
  
ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA  
^  
116 CLAI,  
  
ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
122 CCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATCACGTACTCCACC  
GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG  
  
TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGGAATACTGTATTATTAAACA  
  
AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCTTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAACCTGGTT  
  
AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal  
302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
CGTCTCTGACGCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
^  
303 ALWN1,  
  
ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
362 ACTGTGCCCCATCCCAACATCGAGGAGGTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA  
  
TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGGAGACATCTCATCTTCTGTCAT  
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA  
  
SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
482 TCAAAGAAGAAGTGCGACGAACTCGCCGCAAAGCTGGTGCATTGGGCATCAATGCCGTG  
AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC  
  
AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
542 GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTGTCGTCGTG  
CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTGC CGGCTACAACAGCAGCAC  
^ ^  
550 SAC2, 560 DRD1,  
  
AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
602 GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
^  
615 BSPH1,  
  
ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle

FIG. 14-1



662 ACGTGTGTCACCCAGACAGTCGATTTCAGCCTTGACCTACCTTCACCATTGAGACAATC  
TGCACACAGTGGGTCTGTCTAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTCTGTTAG

ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
722 ACGCTCCCCAAGATGCTGTCTCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG  
TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCTGACCGTCCCCCTTC

ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
782 CCAGGCATCTACAGATTTGTGGCACCAGGGGAGCGCCCCCTCCGGCATGTTCTGACTCGTCC  
GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG

816 BGLI, 833 DRD1,

ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
842 GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCCGCCGAGACT  
CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA

881 SACI,

ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
902 ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTTGGTAGAA

931 SMAI XMAI,

GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
962 GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACCTTCTATCCCAG  
CTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC

985 STUI,

ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
1022 ACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
TGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA

1069 DRA3,

ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
1082 AGGGCTCAAGCCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTGCGCTCAAG  
TCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTTCACAACTAAGCGGAGTTC

ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
1142 CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAAGTGAATC  
GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG

1150 NCOI,

ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
1202 ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC  
TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
1262 GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCTGGCTGCTTTGGCCGCGTATTGCCTG  
CAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATAACGGAC

FIG. 14-2



1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCGGGAAGCCGGCAATCATA  
AGTTGTCCGACGCACCAGTATCACCCGTCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT  
^  
1369 NAEI,

1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA  
GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT  
^  
1385 DRD1,

1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC  
GGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCGTGTGTCCAGACCAACTGGCAA  
GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
^ ^  
1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG  
TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC  
^ ^  
1565 XHOI, 1586 NDEI,

1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
GCGGGCTTGTCACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
CGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAACGAAGTAACCTACCGAAAATGTCTGA  
^ ^  
1643 BSTE2, 1677 ALWN1 PVU2,

1682 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
GCTGTCAACAGCCCACTAACCCTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG  
CGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

1742 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
GTGGCTGCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
CACCGACGGGTCGAGCGGGCGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA  
^  
1794 ESI1,

1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
GGCGCCGCCATCGGCAGTGTTGGACTGGGGAAGGTCCTCATAGACATCCTTGAGGGTAT  
CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCAT  
^  
1802 KAS1 NARI,

1862 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
GGCGCGGGCGTGGCGGGAGCTCTTGTTGGCATTCAAGATCATGAGCGGTGAGGTCCCTCC  
CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
^ ^  
1878 SACI, 1899 BSPH1,

FIG. 14-3





1922 ThrGluAspLeuValAsnLeuLeuPr AlaIleLeuSerProGlyAlaLeuValValGly  
ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
1928 TTH3I,  
ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
1982 GTGGTCTGTGCAGCAATACTGCCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG  
CACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGCTCCCCCGTCACGTCACC  
2004 NAEI, 2017 SMAI XMAI,  
MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
2042 ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCACGCACTACGTG  
TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC  
2067 SMAI XMAI, 2093 DRA3,  
ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
2102 CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC  
2115 PVU2, 2159 ALWN1,  
LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
2162 CTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC  
GAGGACTCCGCTGACGTGGTCACCTATTTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
2164 MST2, 2220 ECON1,  
TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
2222 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT  
LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
2282 AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCTGCCAGCGCGGGTAT  
TTTCGATTTCGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTTCGCGCCCATA  
2285 ESP1, 2300 PVU2, 2310 BAMHI,  
LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
2342 AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
TTCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG  
ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
2402 ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG  
TGACCTGTACAGTTTTTGGCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGATC  
2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,  
TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
2462 TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCCTGTACCCCCCTTCTGCG  
ACCTACCCCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGGAAGGACGC  
2480 ASE1, 2497 APAI,  
ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln

FIG. 14-4



2522 CCGAACTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
GGCTTGATGTGCAAGCGCGATACCTCCACAGACGTCTCCTTATGCACCTCTATTCCGTC  
^

2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCGTGCCAG  
CACCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGCACGGTC  
^

2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
GTCCCATCGCCGAATTTTTTCACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCCC  
CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAACGCGGG

2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTAGAGTAGGACTCCACGAATACCCG  
GGGACGTTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCTGAGGTGCTTATGGGC  
^

2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
GTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC  
CATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACATGCAGGTACGAG  
^

2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGCGAAGGTTGGCGAGGGGATCACCC  
TGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCGCTTCCAACCGCTCCCTAGTGGG  
^

2850 EAG1 XMA3,

2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC  
GGGAGACACCGGTCGAGGAGCCGATCGGTCGATAGGCGAGGTAGAGAGTTCCGTTGAACG  
^

2889 BALI, 2903 NHEI,

2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
ACCGCTAACCATGACTCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
TGCGGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC  
^ ^

2966 ESP1, 2969 SACI,

3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGTTGATTCTGGACTCC  
CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTGTTTACCACCTAAGACCTGAGG

3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC  
^

3096 BGL2,

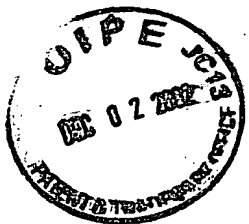
3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgPr AspTyrAsnPro  
CGGAAGTCTCGGAGATTCGCCCAGGCCCTGCCCGTTTGGGCGCGGCCGACTATAACCCC

FIG. 14-5



GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG  
3143 ALWN1, 3164 EAG1 XMA3,  
ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
3182 CCGCTAGTGGAGACGTGGAAAAAGCCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCG  
GGCGATCACCTCTGCACCTTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC  
3217 HGIE2, 3229 NCOI,  
LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
3242 CTTCACCTCCAAAGTCCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTC  
GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG  
ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
3302 ACTGAATCAACCCTATCTACTGCCTTGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC  
TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGCTCTTCGAAACCGTCGAGG  
3332 SACI, 3346 HIND3,  
SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
3362 TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCTTCTGGC  
AGTTGAAGGCCGTAATGCCCCGTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG  
CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
3422 TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT  
ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA  
3437 EAM11051,  
GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
3482 GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG  
CCCCTAGGCCTAGAATCGCTGCCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC  
3484 BAMHI, 3485 BSAB1, 3487 BSPE1,  
AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
3542 GATGTCGTGTGCTGCTCAATGTCTTACTCTTGACAGGCGCACTCGTCACCCCGTGCGCC  
CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTTGGGGCACGCGG  
3589 DRA3, 3600 SAC2,  
AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
3602 GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT  
CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTCTGTTGAGCAACGATGCAGTGGTGTTA  
3611 ALWN1, 3655 PFLM1,  
LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
3662 TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAAGTCACATTTGAC  
AACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGTCTTCTTTTCAGTGTAAGCTG  
3681 DRA3,  
ArgLeuGlnValLeuAspS rHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
3722 AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG  
TCTGACGTTCAAGACCTGTCGGTAATGGTCTGTCATGAGTTCCTCCAATTTCTGTCGCCGC

FIG. 14-6



3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCACAC  
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGTGTG  
3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTCTGCAGGCAACGGTACGGTCTTTCCGG  
3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGTCGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCAGCATTC

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTCATCGTGTTCCTCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCCTTGGCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
AGTGGTCCTGTCGCCCACTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTTGGGGTTAC  
4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC  
4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGCTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTTCGGGCGCACCGGTAGTTCAGG  
4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCAGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC  
GAGTGGCTCTCCGAAATACAACCCCCGGGAGAAATGGTTAAGTCCCCCTCTTGACGCCG  
4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGCGAGCGGCGTACTGACAAGTCTGTTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG

FIG. 14-7



4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyL uGlnAspCysThrMetLeuVal  
TACATCAAGGCCCGGGCAGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGTG  
ATGTAGTTCCGGGCCCGTCGGACAGCTCGGCGTCCCGAGGTCTGACGTGGTACGAGCAC  
^

4452 SMAI XMAI,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGGCGAGC  
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCTCTGCGCCGCTCG  
^ ^

4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCTGGGGACCCCCACAA  
GACTCTCGGAAGTGCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGTT

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCCCACGAC  
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG  
^

4637 SACI,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
GGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCGAGAGCT  
CCGCGACCTTTCTCCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA  
^

4731 NRUI,

4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT  
CGCACCTCTGTCTGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAAA

4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
GCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC  
CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG  
^ ^

4806 PFLM1, 4807 DRA3,

4862 ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu  
AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
TCCCTGGTCGAACCTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT  
^

4893 BGL2,

4922 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG  
^

4954 NCOI,

4982 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAACTTGGGGTACCG  
TCAATGAGAGGTCCACTTTAGTTATCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC  
^ ^

5015 SPHI, 5035 KPNI,

ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly

FIG. 14-8



5042 CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT  
5064 APAI, 5091 BALI,  
GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAA  
CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTTCGAGTTT  
5113 NDEI,  
LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
5162 CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCGGGCTGGTTCACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG  
5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,  
SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
5222 AGCGGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCCGGGGCGACCTAGACCAAACG  
5240 DRA3,  
LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgOP  
5282 CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCAACCGATGAATAGTCGAC  
GATGAGGACGAACGACGTCCCATCCGTAGATGGAGGAGGGGTGGCTACTTATCAGCTG  
5295 PSTI, 5336 SALI,

FIG. 14-9

RECEIVED

DEC 6 1993

FROM GENETIC ENGINEERING

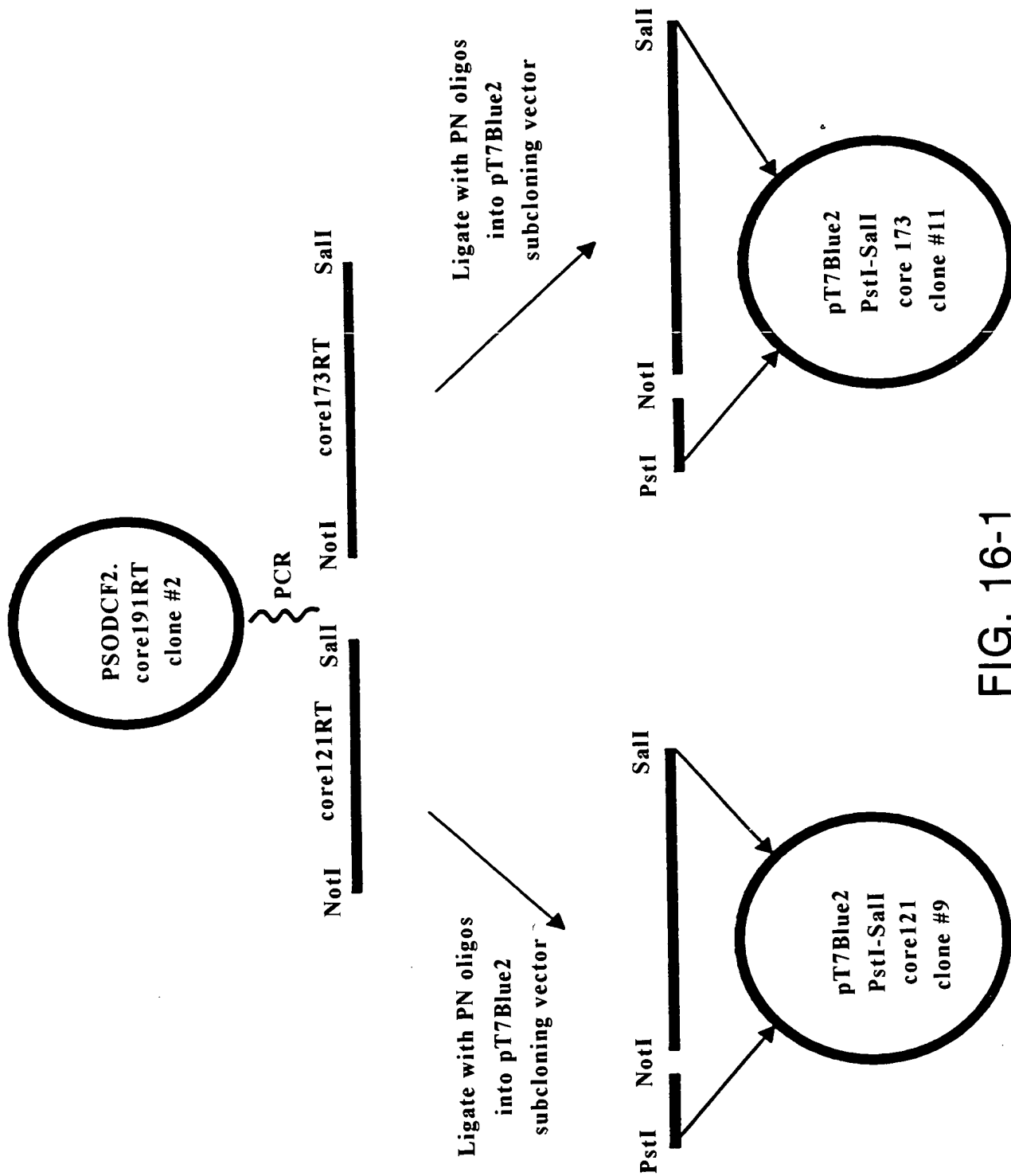
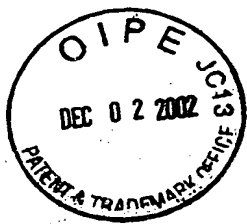
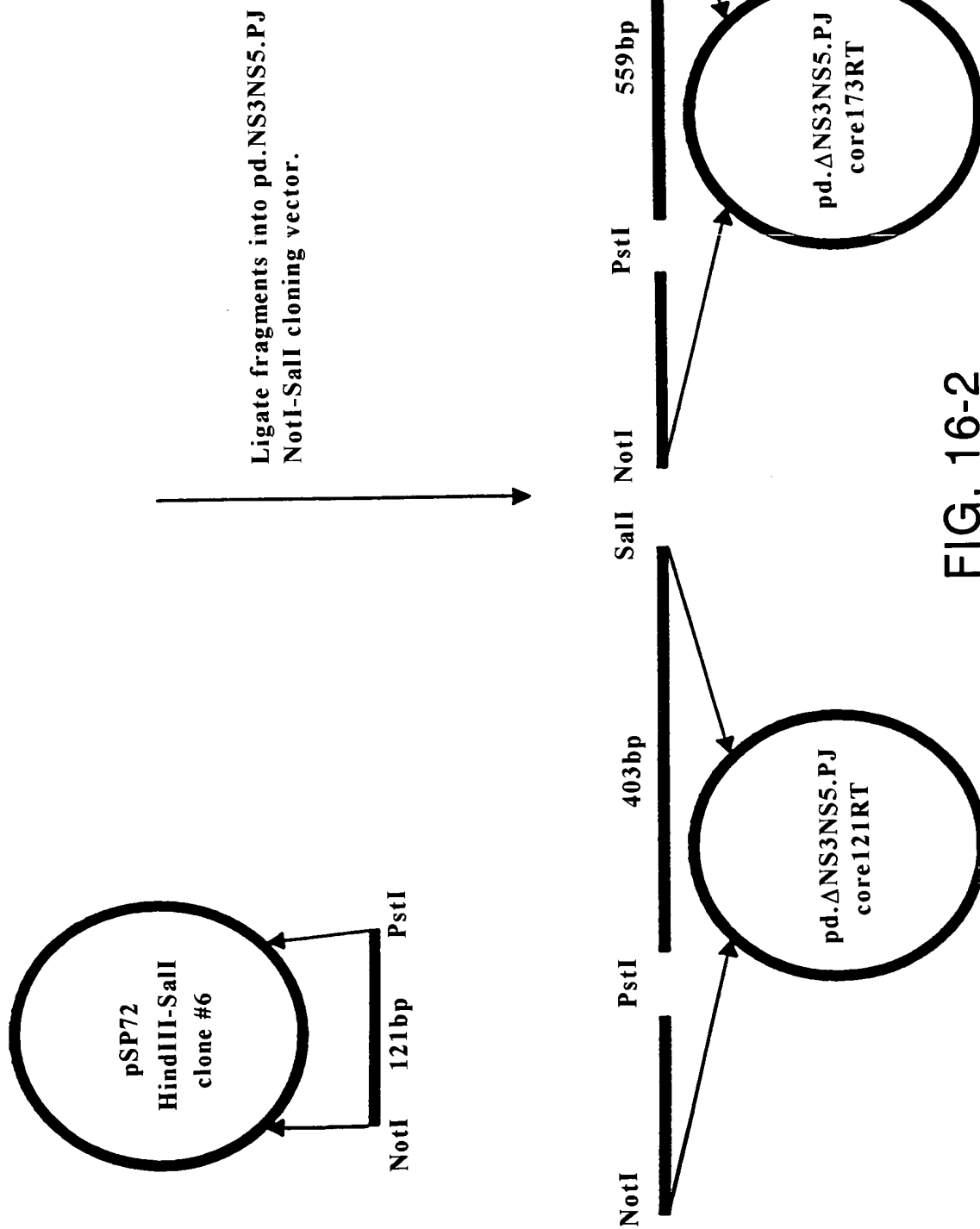
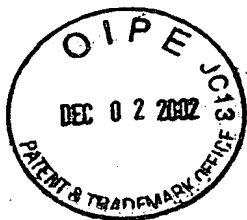


FIG. 16-1







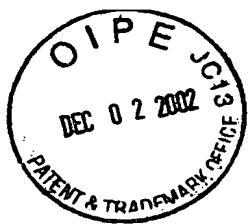
MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
^ ^ ^  
1 HIND3, 24 NDEI, 52 SCAI,  
  
ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA  
^  
116 CLAI,  
  
ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
122 CCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATCACGTACTCCACC  
GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG  
  
TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCCACGAGCCCCCGGAATACTGTATTATTAAACA  
  
AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAAGTGGTT  
  
AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal  
302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
^  
303 ALWN1,  
  
ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA  
  
TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTCAT  
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA

FIG. 17-1

RECEIVED

DEC 9 8 2005

TECH CENTER 18015000



482 SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIle AsnAlaVal  
TCAAAGAAGAAGTGCAGCACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG  
AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC

542 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTGTCGTG  
CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
550 SAC2, 560 DRD1,

602 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
615 BSPH1,

662 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle  
ACGTGTGTCAACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC  
TGCACACAGTGGGTCTGTCTAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTCTGTTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
ACGCTCCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGGAG  
TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCCTGACCGTCCCCCTTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
CCAGGCATCTACAGATTTGTGGCACCAGGGGGAGCGCCCCCTCCGGCATGTTGACTCGTCC  
GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG  
816 BGLI, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCCGAGACT  
CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA  
881 SACI,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTTGTTAGAA  
931 SMAI XMAI,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCCAG  
CTTAAAACCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC  
985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
ACAAAGCAGAGTGGGGAGAACCTTCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
TGTTTCGTCTACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA  
1069 DRA3,

1082 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
AGGGCTCAAGCCCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTGCGCTCAAG

FIG. 17-2



TCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTTCACAACTAAGCGGAGTTC

1142 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTCAGAATGAAATC  
GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG  
^

1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
ACCCTGACGCACCCAGTCACCAATACATCATGACATGCATGTGCGCCGACCTGGAGGTC  
TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG  
^ ^ ^ ^ ^

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTATTGCCTG  
CAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATAACGGAC

1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA  
AGTTGTCCGACGCACCAAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT  
^

1369 NAEI,

1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA  
GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCAGGAGATCGTGAAT  
^

1385 DRD1,

1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC  
GGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA  
GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
^ ^

1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAAGTTCATCAGTGGGATACAATACTTG  
TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC  
^ ^

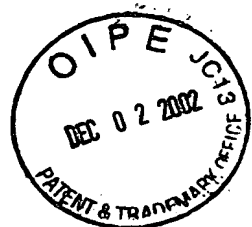
1565 XHOI, 1586 NDEI,

1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
GCGGGCTTGTCACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
CGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAACGAAGTAAGTACCGAAAATGTGCA  
^ ^

1643 BSTE2, 1677 ALWN1 PVU2,

1682 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
GCTGTCAACAGCCCACTAACCCTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGGTGG  
CGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

FIG. 17-3



1742 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
GTGGCTGCCCAGCTCGCCGCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
CACCGACGGGTCGAGCGGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA

1794 ESP1,

1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
GGCGCCGCCATCGGCAGTGTTGGACTGGGGAAGGTCCTCATAGACATCCTTGACAGGGTAT  
CCGCGGCGGTAGCCGTCAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATATA

1802 KAS1 NARI,

1862 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
GGCGCGGGCGTGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC  
CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG

1878 SACI, 1899 BSPH1,

1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG

1928 TTH3I,

1982 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG  
CACCAGACACGTCGTTATGACGCGGGCGTGCAACCGGGCCCGCTCCCCCGTCACGTCACC

2004 NAEI, 2017 SMAI XMAI,

2042 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCACGCACTACGTG  
TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC

2067 SMAI XMAI, 2093 DRA3,

2102 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
CCGGAGAGCGATGCAGCTGCCCCGCTCACTGCCATACTCAGCAGCTCACTGTAACCCAG  
GGCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTCT

2115 PVU2, 2159 ALWN1,

2162 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
CTCCTGAGGCGACTGCACCACTGGATAAGCTCGGAGTGTAACCTCCATGCTCCGGTTCC  
GAGGACTCCGCTGACGTGGTCACTATTTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG

2164 MST2, 2220 ECON1,

2222 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT

2282 LysAlaLysLeuM tProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCTGCCAGCGCGGGTAT  
TTTCGATTTCGAGTACGGTGTTCGACGGACCCTAGGGGAAACACAGGACGGTTCGCGCCCATATA

2285 ESP1, 2300 PVU2, 2310 BAMHI,





2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
ACCGCTAACCATGACTCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC  
^ ^  
2966 ESP1, 2969 SACI,  
3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC  
CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG  
3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC  
^  
3096 BGL2,  
3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro  
CGGAAGTCTCGGAGATTGCCCCAGGCCCTGCCCCGTTTGGGCGCGGCCGACTATAACCCC  
GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG  
^ ^  
3143 ALWN1, 3164 EAG1 XMA3,  
3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCCG  
GGCGATCACCTCTGCACCTTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGG  
^ ^  
3217 HGIE2, 3229 NCOI,  
3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
CTTCCACCTCCAAAGTCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGCTCCTC  
GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG  
3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC  
TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG  
^ ^  
3332 SACI, 3346 HIND3,  
3362 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCCGCCCTTCTGGC  
AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG  
3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT  
ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA  
^  
3437 EAM11051,  
3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
GGGGATCCGGATCTTAGCGACGGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG  
CCCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC  
^ ^ ^  
3484 BAMHI, 3485 BSAB1, 3487 BSPE1,  
3542 AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
GATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC  
CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG

FIG. 17-6



3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT  
CGCCTTCTTGCTTTGACGGGTAGTTACGTGATTTCGTTGAGCAACGATGCAGTGGTGTTA  
3611 ALWN1, 3655 PFLM1,  
3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
TTGGTGATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC  
AACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGTCTTCTTTCAGTGTAAGT  
3681 DRA3,  
3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTAAAGCAGCGGCG  
TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTCGTCGCCG  
3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACGCTGACGCCCCACAC  
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG  
3816 HIND3,  
3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG  
3875 AAT2, 3890 BGLI,  
3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG  
3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGTCTGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCAGCATTC  
4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTCATCGTGTTCCTCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGTTTTCTACCGAAAC  
4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCTTGGCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG  
4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
AGTGGTCTGTGCCCCAATTAAAGGAGCACGTTTCGCACCTTCAGGTTCTTTTGGGGTTAC  
4160 ECORI,  
4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGGCGACGAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC

FIG. 17-7



4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCAAGCCCGCGTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTTCGGGCGCACCGGTAGTTTCAGG

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC  
GAGTGGCTCTCCGAAATACAACCCCGGGAGAATGGTTAAGTTCCCCCTCTTGACGCCG

4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGCAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG

4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
TACATCAAGGCCCCGGGCAGCCTGTGCGAGCCGAGGGCTCCAGGACTGCACCATGCTCGTG  
ATGTAGTTCCGGGCCCGTCCGGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCAC

4452 SMAI XMAI,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGGCGAGC  
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCTCTCTGCGCCGCTCG

4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCTGGGGACCCCCACAA  
GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGT

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCCCCACGAC  
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG

4637 SACI,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
GGCGCTGGAAAGAGGGTCTACTACCTACCCGTGACCCTACAACCCCCCTCGCGAGAGCT  
CCGCGACCTTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA

4731 NRUI,

4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT  
CGCACCTCTGTGCTTCTGTGTGAGGTGAGTTAAGGACCGATCCGTTGTATTAGTACAA

4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
GCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC  
CGGGGGTGAGACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG

4806 PFLM1, 4807 DRA3,

ArgAspGlnLeuGluGlnAlaLeuAspCysGluIle TyrGlyAlaCysTyrSerIle Glu

FIG. 17-8





4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
TCCCTGGTCTGAACCTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT  
^

4893 BGL2,

4922 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG  
^

4954 NCOI,

4982 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTTGGGGTACCG  
TCAATGAGAGGTCCACTTTAGTTATCCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC  
^

5015 SPHI, 5035 KPNI,

5042 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT  
^

5064 APAI, 5091 BALI,

5102 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAA  
CCGTCCCACGCTATACACCGTTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT  
^

5113 NDEI,

5162 LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCGACCGGTTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG  
^^ ^

5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,

5222 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCCGGGGCGACCTAGACCAAACG  
^

5240 DRA3,

5282 LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCAACCGAATGAGCACGAAT  
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTTACTCGTGCTTA  
^

5295 PSTI,

5342 ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
CCTAAACCTCAAAGAAAGACCAAACGTAACACCAACCGCGGCCGAGGACGTCAAGTTC  
GGATTTGGAGTTTCTTTCTGGTTTGCATTGTGGTTGGCCGCCGGCGTCTCTGCAGTTCAAG  
^^ ^

5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,

5402 ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuL uProArgArgGlyProArgLeu  
CCGGGTGGCGGTGAGATCGTTGGTGGAGTTTACTTGTTCGCGCGAGGGGCCCTAGATTG  
GGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC  
^

FIG. 17-9



5449 APAI,

5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTCGCAACCTCGAGGTAGACGTCAGCCT  
CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA  
5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
ATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC  
TAGGGGTTCCGAGCAGCCGGGCTCCCGTCCTGGACCCGAGTCGGGCCCCATGGGAACCGGG  
5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG  
GAGATACCGTTACTCCCGACGCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysOC AM  
CCTAGCTGGGGCCCCACAGACCCCGGCGTAGGTCGCGCAATTTGGGTAAGTAATAGTCG  
GGATCGACCCCGGGGTGTCTGGGGCCGCATCCAGCGCGTTAAACCCATTTCATTATCAGC  
5650 APAI, 5698 SALI,

5702 AC  
TG

FIG. 17-10

RECEIVED

DEC 11 2002

TECH. CENTER, KODAK



MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
^ ^ ^  
1 HIND3, 24 NDEI, 52 SCAI,  
  
ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
62 CCCTCTGTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA  
^  
116 CLAI,  
  
ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
122 CCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATCACGTACTCCACC  
GGATTGTAGTCTTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG  
  
TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
182 TACGGCAAGTTCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCCCACGAGCCCCCGCGAATACTGTATTATTAAACA  
  
AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAAGTGGTT  
  
AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal  
302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
^  
303 ALWN1,  
  
ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA  
  
TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTCAT  
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA  
  
SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
482 TCAAAGAAGAAGTGCGACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG  
AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC  
  
AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
542 GCCTACTACCGCGGTCTTGACGTGTCCGTATCCCGACCAGCGGCGATGTTGTCTGTCGTG  
CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
^ ^ ^  
550 SAC2, 560 DRD1,  
  
AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
602 GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
^  
615 BSPH1,

FIG. 18-1



662 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle  
ACGTGTGTCAACCAGACAGTCGATTTACGCCCTTGACCCTACCTTCACCATTGAGACAATC  
TGCACACAGTGGGTCTGTCAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTCTGTTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
ACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG  
TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCCTGACCGTCCCCCTTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
CCAGGCATCTACAGATTTGTGGCACCAGGGGAGCGCCCTCCGGCATGTTTCGACTCGTCC  
GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG  
^ ^

816 BGLI, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCCGAGACT  
CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA  
^

881 SACI,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTGGTAGAA  
^

931 SMAI XMAI,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCACTTTCTATCCAG  
CTTAAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC  
^

985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
ACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
TGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA  
^

1069 DRA3,

1082 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
AGGGCTCAAGCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTGCGCTCAAG  
TCCCGAGTTCGGGGAGGGGGTAGCACCCCTGGTCTACACCTTCACAACTAAGCGGAGTTC

1142 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCAAGATGAAATC  
GGGTGGGAGGTACCCGGTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG  
^

1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC  
TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG  
^ ^ ^ ^ ^

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrS rThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCTGGCTGCTTTGGCCGCGTATTGCCTG



CAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATAACGGAC

1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGAAGCCGGCAATCATA  
AGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT  
^

1369 NAEI,

1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGTCTCTCAGCACTTA  
GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT  
^

1385 DRD1,

1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC  
GGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
CTGCAGACCGCGTCCCGTCAGGCAGAGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA  
GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
^ ^

1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG  
TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC  
^ ^

1565 XHOI, 1586 NDEI,

1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
CGCCCGAACAGTTGCGACGGACCATTTGGGGCGGTAACGAAGTAACTACCGAAAATGTCA  
^ ^

1643 BSTE2, 1677 ALWN1 PVU2,

1682 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
GCTGTCACCAGCCCACTAACCCTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG  
CGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

1742 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
GTGGCTGCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
CACCGACGGGTGAGCGGGCGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA  
^

1794 ESI1,

1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
GGCGCCGCCATCGGCAGTGTGGACTGGGGAAGGTCCTCATAGACATCCTTGCAGGGTAT  
CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCAT  
^

1802 KAS1 NARI,

1862 GlyAlaGlyValAlaGlyAlaL uValAlaPheLysIleMetSerGlyGluValProSer  
GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCTCC  
CCGCGCCCGCACCCGCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
^ ^

1878 SACI, 1899 BSPH1,

FIG. 18-3



1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
TGCCTCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
1928 TTH3I,  
ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
1982 GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCACTGG  
CACCAGACACGTCGTTATGACGCGGGCGTGCAACCGGGCCCGCTCCCCCGTCACGTCACC  
2004 NAEI, 2017 SMAI XMAI,  
MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
2042 ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGAACCATGTTTCCCCACGCACTACGTG  
TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC  
2067 SMAI XMAI, 2093 DRA3,  
ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
2102 CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTGCGAGTGACATTGGGTC  
2115 PVU2, 2159 ALWN1,  
LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
2162 CTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC  
GAGGACTCCGCTGACGTGGTCACCTATTTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
2164 MST2, 2220 ECON1,  
TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
2222 TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT  
LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
2282 AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCCTTTGTGTCTGCCAGCGCGGGTAT  
TTTCGATTGAGTACGGTGTGACGGACCCTAGGGGAAACACAGGACGGTCGCGCCCATA  
2285 ESP1, 2300 PVU2, 2310 BAMHI,  
LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
2342 AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
TTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG  
ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
2402 ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG  
TGACCTGTACAGTTTTTGGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGTA  
2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,  
TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
2462 TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGGCCCCTGTACCCCCCTTCCTGCG  
ACCTCACCTGGAAGGGGTAAATTACGGATGTGGTGCCCGGGGACATGGGGGGAAGGACGC  
2480 ASE1, 2497 APAI,

FIG. 18-4



2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln  
CCGAACACTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
GGCTTGATGTGCAAGCGCGATACCTCCCACAGACGTCTCCTTATGCACCTCTATTCCGTC  
^

2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCCGTGCCAG  
CACCCCTGAAGGTGATGCACTGCCCATCTGATGACTGTTAGAATTTACGGGCACGGTC  
^

2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
GTCCCATCGCCGAATTTTTTCACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCCC  
CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAACGCGGG

2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTGAGAGTAGGACTCCACGAATACCCG  
GGGACGTTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC  
^

2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
GTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC  
CATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACGAGGTACGAG  
^

2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGGCGAAGGTTGGCGAGGGGATCACCC  
TGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCCGCTTCCAACCGCTCCCCTAGTGGG  
^

2850 EAG1 XMA3,

2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC  
GGGAGACACCGGTCGAGGAGCCGATCGGTTCGATAGGCGAGGTAGAGAGTTCCGTTGAACG  
^

2889 BALI, 2903 NHEI,

2942 ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
TGCGGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC  
^ ^

2966 ESP1, 2969 SACI,

3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC  
CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG

3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC  
^

3096 BGL2,

ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro

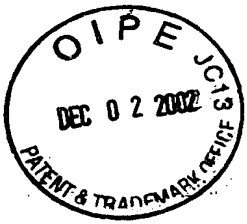
FIG. 18-5



3122 CGGAAGTCTCGGAGATTGCGCCAGGCCCTGCCCGTTTGGGCGCGGCCGGACTATAACCCC  
GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG  
^ ^  
3143 ALWN1, 3164 EAG1 XMA3,  
ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
3182 CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACACCTGTGGTCCATGGCTGCCCG  
GGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC  
^ ^  
3217 HGIE2, 3229 NCOI,  
LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
3242 CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCCTCGGAAGAAGCGGACGGTGGTCCTC  
GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG  
ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
3302 ACTGAATCAACCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC  
TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG  
^ ^  
3332 SACI, 3346 HIND3,  
SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
3362 TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCTTCTGGC  
AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG  
CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
3422 TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT  
ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA  
^ ^  
3437 EAM11051,  
GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
3482 GGGGATCCGGATCTTAGCGACGGGTTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG  
CCCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC  
^ ^ ^  
3484 BAMHI, 3485 BSAB1, 3487 BSPE1,  
AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
3542 GATGTGCTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC  
CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG  
^ ^  
3589 DRA3, 3600 SAC2,  
AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
3602 GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTACCCACAAT  
CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTGCTTGAGCAACGATGCAGTGGTGTTA  
^ ^  
3611 ALWN1, 3655 PFLM1,  
LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
3662 TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC  
AACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGTCTTCTTTCAAGTGTAAACTG  
^ ^  
3681 DRA3,  
ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
3722 AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG

FIG. 18-6





TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTTCTGTCGCCGC

3782 SerLysValLysAlaAsnLeuLeuS rValGluGluAlaCysSerLeuThrProProHis  
TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACGCTGACGCCCCACAC  
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG

3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGTCTGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCAGCATTC

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTCATCGTGTTCCTCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCTTGCGCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
AGTGGTCTGTGCGCCCACTTAAGGAGCACGTTTCGCACCTTCAGGTTCTTTTGGGGTTAC

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGGCGACGAAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC

4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGGTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTGCGGCGCACCGGTAGTTCAGG

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCAGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC  
GAGTGGCTCTCCGAAATACAACCCCGGGAGAATGGTTAAGTTCCCCCTCTTGACGCCG

4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGGAGCGGCGTACTGACAAGTCTGTTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG

FIG. 18-7



4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
TACATCAAGGCCCGGGCAGCTTGTGAGCCGCGAGGGCTCCAGGACTGCACCATGCTCGTG  
ATGTAGTTCCGGGGCCCGTCGGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCAC  
^  
4452 SMAI XMAI,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGGCGAGC  
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCCTGCGCCGCTCG  
^ ^  
4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCTGGGGACCCCCACAA  
GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGT  
^

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTGAGTCGCCCACGAC  
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG  
^  
4637 SACI,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
GGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCGAGAGCT  
CCGCGACCTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA  
^  
4731 NRUI,

4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT  
CGCACCTCTGTGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAAA

4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
GCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC  
CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG  
^ ^  
4806 PFLM1, 4807 DRA3,

4862 ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu  
AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
TCCCTGGTCAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGACGATGAGGTATCTT  
^  
4893 BGL2,

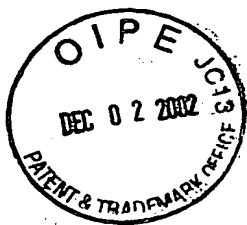
4922 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAGTGAGGTG  
^  
4954 NCOI,

4982 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAACTTGGGGTACCG  
TCAATGAGAGGTCCACTTTAGTTATCCCACCGGTACGGAGTCTTTTGAACCCCATGGC  
^ ^  
5015 SPHI, 5035 KPNI,

FIG. 18-8



5042 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT  
5064 APAI, 5091 BALI,  
5102 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAA  
CCGTCCCGACGGTATACACCGTTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT  
5113 NDEI,  
5162 LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCGACCGGTTCGACCTGAACAGGCGGACCAAGTGCCGACCGATG  
5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,  
5222 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCGGGGCGACCTAGACCAAACG  
5240 DRA3,  
5282 LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCAACCGAATGAGCACGAAT  
GATGAGGACGAACGACGTCCCATCCGTAGATGGAGGAGGGGTTGGCTTACTCGTGCTTA  
5295 PSTI,  
5342 ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
CCTAAACCTCAAAGAAAGACCAAACGTAACACCAACCGGCGGCCGAGGACGTCAAGTTC  
GGATTTGGAGTTTCTTTCTGGTTTGCATTGTGGTTGGCCGCCGCGTCTCTGCAGTTCAAG  
5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,  
5402 ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
CCGGGTGGCGGTCAGATCGTTGGTGGAGTTTACTTGTGTCGCGCAGGGGCCCTAGATTG  
GGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC  
5449 APAI,  
5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTCGCAACCTCGAGGTAGACGTCAGCCT  
CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA  
5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,  
5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
ATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC  
TAGGGGTTCCGAGCAGCCGGGCTCCCGTCTGGACCCGAGTCGGGCCCATGGGAACCGGG  
5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,  
5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCGTGGCTCTCGG  
GAGATACCGTTACTCCCGACGCCCCACCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC



ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp  
5642 CCTAGCTGGGGCCCCACAGACCCCGGCGTAGGTCGCGCAATTTGGGTAAGGTCATCGAT  
GGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCATTCCAGTAGCTA  
5650 APAI, 5696 CLAI,  
ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValGlyAlaProLeu  
5702 ACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTCGTCGGCGCCCCCTCTT  
TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGCCGCGGGGAGAA  
5724 HGIE2, 5750 KAS1 NARI, 5756 ECON1,  
GlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgValLeuGluAspGlyValAsnTyr  
5762 GGAGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTCTGGAAGACGGCGTGAACTAT  
CCTCCGCGACGGTCCCGGGACCGCGTACCGCAGGCCCAAGACCTTCTGCCGCACTTGATA  
5772 BSTXI, 5775 APAI,  
AlaThrGlyAsnLeuProGlyCysSerOC AM  
5822 GCAACAGGGAACCTTCCTGGTTGCTCTTAATAGTCGAC  
CGTTGTCCCTTGGAAGGACCAACGAGAATTATCAGCTG  
5854 SALI,

FIG. 18-10

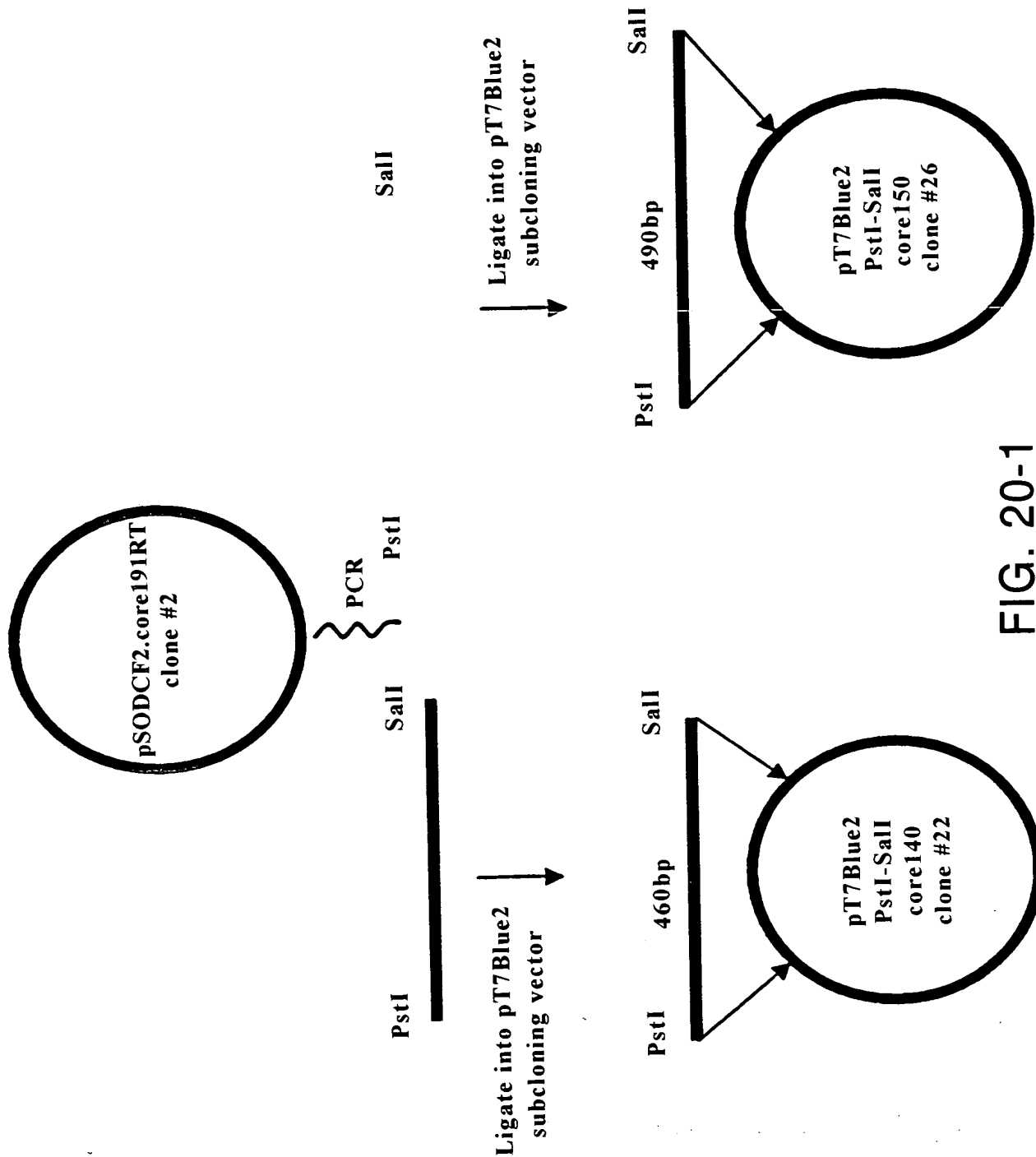
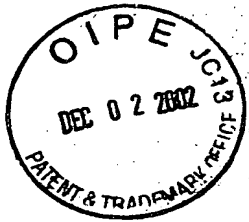


FIG. 20-1

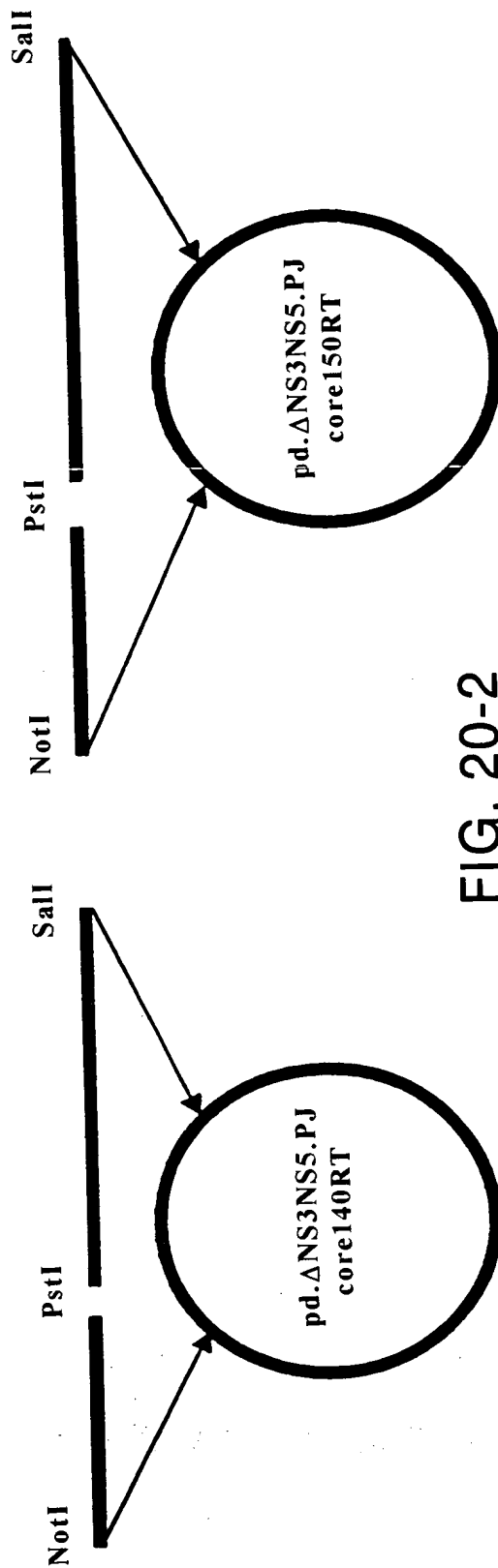
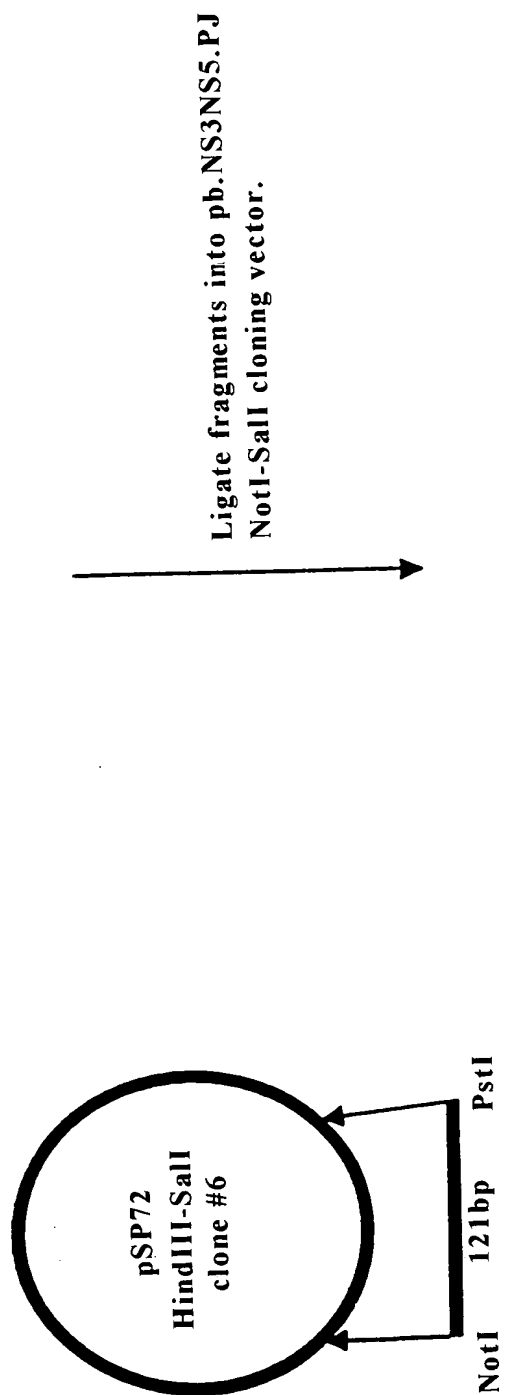
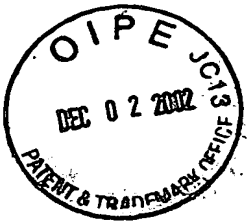


FIG. 20-2



MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
1 HIND3, 24 NDEI, 52 SCAI,  
ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA  
116 CLAI,  
ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
122 CCTAACATCAGGACCGGGGTGAGAACAAATTACCACTGGCAGCCCCATCACGTACTCCACC  
GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGATGAGGTGG  
TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGCGAATACTGTATTATTAAACA  
AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAAGTGGTT  
AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal  
302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
303 ALWN1,  
ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA  
TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTCAT  
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA

FIG. 21-1

RECEIVED  
DEC 02 2002  
DEPT. OF COMMERCE



482 SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
TCAAAGAAGAAGTGCACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG  
AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC

542 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTCTCGTCGTG  
CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
550 SAC2, 560 DRD1,

602 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
615 BSPH1,

662 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle  
ACGTGTGTCACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC  
TGCACACAGTGGGTCTGTCAGCTAAAGTCGGAACCTGGGATGGAAGTGGAACCTCTGTTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
ACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG  
TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCCTGACCGTCCCCCTTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
CCAGGCATCTACAGATTGTGGCACCAGGGGAGCGCCCTCCGGCATGTTGACTCGTCC  
GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG  
816 BGL1, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCCGAGACT  
CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA  
881 SAC1,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTGTGTAGAA  
931 SMAI XMAI,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCAG  
CTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC  
985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
ACAAAGCAGAGTGGGGAGAACCTTCCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
TGTTTCGTCTACCCCTCTTGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA  
1069 DRA3,

1082 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysL uIleArgLeuLys  
AGGGCTCAAGCCCCTCCCCATCGTGGGACCAGATGTGGAAGTGTTTGATTGCGCTCAAG

FIG. 21-2





TCCCGAGTTCGGGGAGGGGGTAGCACCTGGTCTACACCTTCACAACTAAGCGGAGTTC

1142 ProThrLeuHisGlyPr ThrPr LeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle  
CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCTCAGAATGAAATC  
GGGTGGGAGGTACCCGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG  
^

1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC  
TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG  
^ ^ ^ ^ ^

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCTGGCTGCTTTGGCCGCGTATTGCCTG  
CAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATAACGGAC

1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
TCAACAGGCTGCGTGGTTCATAGTGGGCAGGGTCGTCTTGTCGGGAAGCCGGCAATCATA  
AGTTGTCCGACGCACAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT  
^

1369 NAEI,

1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
CCTGACAGGGAAGTCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA  
GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCAGGAGATCGTGAAT  
^

1385 DRD1,

1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC  
GGCATGTAGCTCGTTCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCGGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
CTGCAGACCGCGTCCCGTCAGGCAGAGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA  
GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT  
^ ^

1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG  
TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC  
^ ^

1565 XHOI, 1586 NDEI,

1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
CGCCCGAACAGTTGCGACGGACCATTTGGGGCGGTAACGAAGTAACCTACCGAAAATGTGCA  
^ ^

1643 BSTE2, 1677 ALWN1 PVU2,

1682 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
GCTGTCCAGCCCACTAACCCTAGCCAAACCCTCCTCTCAACATATTGGGGGGGTGG  
CGACAGTGGTGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

FIG. 21-3



1742 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla  
GTGGCTGCCAGCTCGCCGCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
CACCGACGGGTCGAGCGGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA  
1794 ESP1,

1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
GGCGCCGCCATCGGCAGTGTGGACTGGGGAAGGTCCTCATAGACATCCTTGCCAGGGTAT  
CCGCGGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCAT  
1802 KAS1 NARI,

1862 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC  
CCGCGCCCCGACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
1878 SACI, 1899 BSPH1,

1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
1928 TTH3I,

1982 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG  
CACCAGACACGTCGTTATGACGCGGGCGGTGCAACCGGGCCCGCTCCCCCGTCACGTACCC  
2004 NAEI, 2017 SMAI XMAI,

2042 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGGAACCATGTTCCCCACGCACTACGTG  
TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC  
2067 SMAI XMAI, 2093 DRA3,

2102 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
CCGGAGAGCGATGCAGCTGCCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
GGCCTCTCGCTACGTGACGGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC  
2115 PVU2, 2159 ALWN1,

2162 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
CTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTAACCACTCCATGCTCCGGTTCC  
GAGGACTCCGCTGACGTGGTCACCTATTGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
2164 MST2, 2220 ECON1,

2222 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT

2282 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCCTGCCAGCGCGGGTAT  
TTTCGATTTCGAGTACGGTGTGACGGACCCATTAGGGGAAACACAGGACGGTTCGCGCCCAT  
2285 ESP1, 2300 PVU2, 2310 BAMHI,

FIG. 21-4



2342 LysGlyValTrpArgGlyAspGlyIleM tHisThrArgCysHisCysGlyAlaGluIle  
AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
TTCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG

2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG  
TGACCTGTACAGTTTTTGGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGATC  
^ ^ ^  
2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

2462 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCTGTACCCCTTCTGCG  
ACCTCACCTGGAAGGGTAATTACGGATGTGGTGCCCGGGACATGGGGGAAGGACGC  
^ ^  
2480 ASE1, 2497 APAI,

2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln  
CCGAACCTACACGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
GGCTTGATGTGCAAGCGCGATACCTCCACAGACGTCTCCTTATGCACCTCTATTCCGTC  
^  
2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCCGTGCCAG  
CACCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGCACGGTC  
^  
2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
GTCCCATCGCCGAATTTTTACAGAATTGGACGGGGTGGCCTACATAGGTTTGCGCC  
CAGGGTAGCGGGCTTAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCAAACGCGGG

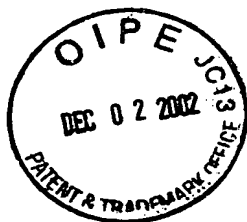
2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTCAGAGTAGGACTCCACGAATACCCG  
GGGACGTTGCGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC  
^  
2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
GTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC  
CATCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACTGCAGGTACGAG  
^  
2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGCGGAAGGTTGGCGAGGGGATCACCC  
TGACTAGGGAGGGTATATTGTCGTCTCCGCCGCGCCGCTTCCAACCGCTCCCTAGTGGG  
^  
2850 EAG1 XMA3,

2882 ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys  
CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC  
GGGAGACACCGGTCGAGGAGCGATCGGTGATAGGCGAGGTAGAGAGTTCCGTTGAACG  
^ ^  
2889 BALI, 2903 NHEI,

FIG. 21-5



2942 ThrAlaAsnHisAspSerPr AspAlaGluLeuIl GluAlaAsnLeuL uTrpArgGln  
ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC  
^ ^  
2966 ESP1, 2969 SACI,

3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC  
CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG

3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC  
^  
3096 BGL2,

3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro  
CGGAAGTCTCGGAGATTCGCCAGGCCCTGCCGTTTGGGCGCGGCCGACTATAACCCC  
GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGSCCTGATATTGGGG  
^ ^  
3143 ALWN1, 3164 EAG1 XMA3,

3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCCG  
GGCGATCACCTCTGCACCTTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC  
^ ^  
3217 HGIE2, 3229 NCOI,

3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCTCGGAAGAAGCGGACGGTGGTCCTC  
GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCTGCCACCAGGAG

3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC  
TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG  
^ ^  
3332 SACI, 3346 HIND3,

3362 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
TCAACTTCGGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCTTCTGGC  
AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG

3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT  
ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA  
^  
3437 EAM11051,

3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG  
CCCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC  
^ ^ ^  
3484 BAMHI, 3485 BSAB1, 3487 BSPE1,

3542 AspValValCysCysSerMetSerTyrS rTrpThrGlyAlaLeuValThrProCysAla  
GATGTCGTGTGCTGCTCAATGTCTTACTCTTGACAGGCGCACTCGTCACCCCGTGCGCC  
CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG

FIG. 21-6



3589 DRA3, 3600 SAC2,

3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT  
CGCCTTCTTGCTTTGACGGGTAGTTACGTGATTCGTTGAGCAACGATGCAGTGGTGTTA

3611 ALWN1, 3655 PFLM1,

3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC  
AACCACATAAGGTGGTGGAGTGCCTCACGAACGGTTTCCGTCTTCTTTCAGTGAAACTG

3681 DRA3,

3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTAAAGCAGCGGCG  
TCTGACGTTCAAGACCTGTCCGTAATGGTCCTGCATGAGTTCCTCCAATTTTCGTCGCCCG

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACGCTGACGCCCCACAC  
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG

3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGTCGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCAGCATTC

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTCATCGTGTTCCTCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGTTTTCTACCGAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCTTGCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
AGTGGTCTGTGCGCCAACTTAAGGAGCACGTTGCGACCTTCAGGTTCTTTTGGGGTTAC

4160 ECORI,

4202 GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCCAAGAGCATACTATGGGCGACGAACTGAGGTGTGAGTACTCTCGCTGTAGGCATGC

FIG. 21-7



4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGGTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCTGGGCGCACCGGTAGTTCAGG

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC  
GAGTGGCTCTCCGAAATACAACCCCGGAGAAATGGTTAAGTTCCCCCTCTTGACGCCG

4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGCGAGCGGCGTACTGACAACCTAGCTGTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG

4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
TACATCAAGGCCCGGGCAGCCTGTGAGCCGCGAGGGCTCCAGGACTGCACCATGCTCGTG  
ATGTAGTTCCGGGCCCCGTCGGACAGCTCGGCGTCCCCGAGGTCTGACGTGGTACGAGCAC

4452 SMAI XMAI,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGGCGAGC  
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCTCTCTGCGCCGCTCG

4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCTGGGGACCCCCACAA  
GACTCTCGGAAGTGCTCCGATACTGCTCCATGAGGCGGGGGGACCCCTGGGGGGTGT

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCTGAGTCGCCCCACGAC  
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGTCTG

4637 SACI,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
GGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCGAGAGCT  
CCGCGACCTTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA

4731 NRUI,

4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT  
CGCACCTCTGTCTGTTCTGTGTGAGGTCTAGTTAAGGACCGATCCGTTGTATTAGTACAAA

4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
GCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCTTTATAGCC  
CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG

4806 PFLM1, 4807 DRA3,

ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu

FIG. 21-8



4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
TCCCTGGTCTGAACCTTGTCGGGAGCTAACGCTCTAGATGCCCCGACGATGAGGTATCTT  
^

4893 BGL2,

4922 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG  
^

4954 NCOI,

4982 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTTGGGGTACCG  
TCAATGAGAGGTCCACTTTAGTTATCCACCGCGTACGGAGTCTTTTGAACCCCATGGC  
^

5015 SPHI, 5035 KPNI,

5042 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCT  
^

5064 APAI, 5091 BALI,

5102 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAA  
CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT  
^

5113 NDEI,

5162 LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCGGGCTGGTTCACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCGACCGGTGACCTGAACAGGCCGACCAAGTGCCGACCGATG  
^

5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,

5222 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGGCCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCGACAGAGTACGGGCCGGGGCGACCTAGACCAAACG  
^

5240 DRA3,

5282 LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACCGAATGAGCACGAAT  
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTTACTCGTGCTTA  
^

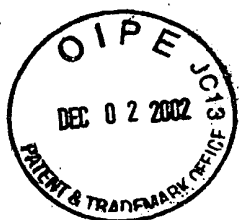
5295 PSTI,

5342 ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
CCTAAACCTCAAAGAAAGACCAAACGTAACACCAACCGGCGGGCCGAGGACGTCAAGTTC  
GGATTTGGAGTTTCTTTCTGGTTTGCAATTGTGGTTGGCCGCGGCGTCTGCAAGTTCAAG  
^

5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,

5402 ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
CCGGGTGGCGGTGAGATCGTTGGTGGAGTTTACTTGTGCGCGCAGGGGGCCCTAGATTG  
GGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC  
^

FIG. 21-9



5449 APAI,

5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTGCAACCTCGAGGTAGACGTCAGCCT  
CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA  
^ ^ ^ ^

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
ATCCCCAAGGCTCGTCGGCCCCGAGGGCAGGACCTGGGCTCAGCCCCGGGTACCCTTGGCCC  
TAGGGGTTCCGAGCAGCCGGGCTCCCGTCCTGGACCCGAGTCGGGCCCATGGGAACCGGG  
^ ^ ^ ^

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG  
GAGATACCGTTACTCCCGACGCCCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp  
CCTAGCTGGGGCCCCACAGACCCCCGGCGTAGGTGCGCAATTTGGGTAAGGTCATCGAT  
GGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGGTTAAACCCATTCCAGTAGCTA  
^ ^

5650 APAI, 5696 CLAI,

5702 ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValOC AM  
ACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTCGTCTAATAGTCGAC  
TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGATTATCAGCTG  
^ ^

5724 HGIE2, 5755 SALI,

FIG. 21-10





MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn  
2 AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC  
TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG  
1 HIND3, 24 NDEI, 52 SCAI,  
ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp  
62 CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT  
GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCCTAGCTA  
116 CLAI,  
ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr  
122 CCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCAGTACTCCACC  
GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG  
TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys  
182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT  
ATGCCGTTCAAGGAACGGCTGCCGCCACGAGCCCCCGCGAATACTGTATTATTAAACA  
AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln  
242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA  
CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAAGTGGTT  
AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal  
302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC  
CGTCTCTGACGCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG  
303 ALWN1,  
ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe  
362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT  
TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA  
TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis  
422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGAGACATCTCATCTTCTGTGCAT  
ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCTCTGTAGAGTAGAAGACAGTA

FIG. 22-1



482 SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal  
TCAAAGAAGAAGTGCGACGAAGCTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG  
AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC

542 AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal  
GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTGTCGTCGTG  
CGGATGATGGCGCCAGAAGTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC  
550 SAC2, 560 DRD1,

602 AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn  
GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT  
CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA  
615 BSPH1,

662 ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle  
ACGTGTGTCACCCAGACAGTCGATTTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC  
TGCACACAGTGGGTCTGTCTCAGCTAAAGTCGGAAGTGGGATGGAAGTGGTAACTCTGTTAG

722 ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys  
ACGCTCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG  
TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCCTGACCGTCCCCCTTC

782 ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer  
CCAGGCATCTACAGATTTGTGGCACCAGGGGAGCGCCCCCTCCGGCATGTTGACTCGTCC  
GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG  
816 BGLI, 833 DRD1,

842 ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr  
GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCCGCGAGACT  
CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA  
881 SACI,

902 ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu  
ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT  
TGTCATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCTGGTAGAA  
931 SMAI XMAI,

962 GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln  
GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCCAG  
CTTAAACCCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC  
985 STUI,

1022 ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla  
ACAAAGCAGAGTGGGGAGAACCTTCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT  
TGTTTCGTCTACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA  
1069 DRA3,

1082 ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys  
AGGGCTCAAGCCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTGTTGATTGCGCTCAAG

FIG. 22-2



TCCCGAGTTCGGGGAGGGGGTAGCACCCCTGGTCTACACCTTCACAACTAAGCGGAGTTC

1142 ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluI  
CCCACCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTTCTGAGATGAAATC  
GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG

1150 NCOI,

1202 ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal  
ACCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTGCGCCGACCTGGAGGTC  
TGGGACTGCGTGGGTGAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

1262 ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu  
GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTATTGCCTG  
CAGTGCTCGTGGACCCACGAGCAACCGCCGAGGACCGACGAAACCGGCGCATAACGGAC

1322 SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle  
TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA  
AGTTGTCCGACGCACCAAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT

1369 NAEI,

1382 ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu  
CCTGACAGGGAAGTCTCTACCGAGAGTTCGATGAGATGGAAGAGTGTCTCTCAGCACTTA  
GGACTGTCCCTTCAGGAGATGGCTCTCAAGTACTCTACCTTCTCAGAGAGTCGTGAAT

1385 DRD1,

1442 ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu  
CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC  
GGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG

1502 LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln  
CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA  
GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT

1502 PSTI, 1507 TTH3I,

1562 LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu  
AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACCTTCATCAGTGGGATACAATACTTG  
TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC

1565 XHOI, 1586 NDEI,

1622 AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla  
GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT  
CGCCCGAACAGTTGCGACGGACCATTTGGGGCGGTAACGAAGTAACCTACCGAAAATGTCTGA

1643 BSTE2, 1677 ALWN1 PVU2,

1682 AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp  
GCTGTCAACAGCCCACTAACCCTAGCCAAACCTCCTCTTCAACATATTGGGGGGGGTGG  
CGACAGTGGTGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCACC

FIG. 22-3



1742 ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPh ValGlyAlaGlyLeuAla  
GTGGCTGCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT  
CACCGACGGGTCGAGCGGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA  
1794 ESP1,

1802 GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr  
GGCGCCGCCATCGGCAGTGTGGACTGGGGAAGGTCCTCATAGACATCCTTGCAGGGTAT  
CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATA  
1802 KAS1 NARI,

1862 GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer  
GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC  
CCGCGCCCGCACCCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG  
1878 SACI, 1899 BSPH1,

1922 ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly  
ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC  
TGCTCTCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG  
1928 TTH3I,

1982 ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp  
GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAAGTGG  
CACCAGACACGTCGTTATGACGCGGGCGTGAACCGGGCCGCTCCCCCGTCACGTCACC  
2004 NAEI, 2017 SMAI XMAI,

2042 MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal  
ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGGAACCATGTTTCCCCCAGCCTACGTCG  
TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC  
2067 SMAI XMAI, 2093 DRA3,

2102 ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln  
CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG  
GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC  
2115 PVU2, 2159 ALWN1,

2162 LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer  
CTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTAACCTCCATGCTCCGGTTCC  
GAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG  
2164 MST2, 2220 ECON1,

2222 TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu  
TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA  
ACCGATTCCCTGTAGACCTGACCTATACGCTCCACAACCTCGCTGAAATTCTGGACCGAT

2282 LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr  
AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCCTTTGTGTCTGCGCAGCGGGGTAT  
TTTCGATTGAGTACGGTGTGACGGACCCCTAGGGGAAACACAGGACGGTCGCGCCCATA  
2285 ESP1, 2300 PVU2, 2310 BAMHI,

FIG. 22-4



2342 LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle  
AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC  
TTCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG

2402 ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet  
ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG  
TGACCTGTACAGTTTTTGGCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGAC  
^ ^ ^

2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

2462 TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla  
TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCTGTACCCCCCTTCCTGCG  
ACCTCACCTGGAAGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGAAGGACGC  
^

2480 ASE1, 2497 APAI,

2522 ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln  
CCGAACCTACAGTTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG  
GGCTTGATGTGCAAGCGCGATACCTCCACAGACGTCCTCTTATGCACCTCTATTCCGTC  
^

2553 PSTI,

2582 ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln  
GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCGTGCCAG  
CACCCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGCACGGTC  
^

2594 DRA3,

2642 ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro  
GTCCCATCGCCGAATTTTTTACAGAATTGGACGGGGTGCCTACATAGGTTTGCGCCC  
CAGGGTAGCGGGCTTAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCAAACGCGGG

2702 ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro  
CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTAGAGTAGGACTCCACGAATACCCG  
GGGACGTTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC  
^

2757 HGIE2,

2762 ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu  
GTAGGGTCGCAATTACCTTGCGAGCCCGAACCAGGACGTGGCCGTGTTGACGTCCATGCTC  
CATCCAGCGTTAATGGAACGCTCGGGCTTGCCCTGCACCGGCACAATGCAGGTACGAG  
^

2809 AAT2,

2822 ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro  
ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGCGAAGGTTGGCGAGGGGATCACCC  
TGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCGCTTCCAACCGCTCCCTAGTGGG  
^

2850 EAG1 XMA3,

2882 ProSerValAlaSerSerS rAlaSerGlnLeuSerAlaProSerL uLysAlaThrCys  
CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC  
GGGAGACACCGGTCGAGGAGCCGATCGGTCGATAGGCGAGGTAGAGAGTTCCGTTGAACG  
^ ^

2889 BALI, 2903 NHEI,

FIG. 22-5



2942 ThrAlaAsnHisAspSerPr AspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln  
ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG  
TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC  
^ ^  
2966 ESP1, 2969 SACI,  
  
3002 GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer  
GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC  
CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTCACCACTAAGACCTGAGG  
  
3062 PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu  
TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG  
AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC  
^  
3096 BGL2,  
  
3122 ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro  
CGGAAGTCTCGGAGATTCGCCCAGGCCCTGCCCGTTTGGGCGCGGCCGACTATAACCCC  
GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG  
^ ^  
3143 ALWN1, 3164 EAG1 XMA3,  
  
3182 ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro  
CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCCG  
GGCGATCACCTCTGCACCTTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGG  
^ ^  
3217 HGIE2, 3229 NCOI,  
  
3242 LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu  
CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTC  
GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG  
  
3302 ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer  
ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC  
TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG  
^ ^  
3332 SACI, 3346 HIND3,  
  
3362 SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly  
TCAACTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCCGCCCTTCTGGC  
AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG  
  
3422 CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro  
TGCCCCCCCCGACTCCGACGCTGAGTCCTATTCTCCATGCCCCCCTGGAGGGGGAGCCT  
ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA  
^  
3437 EAM11051,  
  
3482 GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu  
GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG  
CCCCTAGGCCTAGAATCGCTGCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC  
^ ^ ^  
3484 BAMHI, 3485 BSAB1, 3487 BSPE1,  
  
3542 AspValValCysCysS rMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla  
GATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC  
CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG

FIG. 22-6



3589 DRA3, 3600 SAC2,

3602 AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn  
GCGGAAGAACAGAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT  
CGCCTTCTTGCTTTGACGGGTAGTTACGTGATTCTGTTGAGCAACGATGCAGTGGTGTTA

3611 ALWN1, 3655 PFLM1,

3662 LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp  
TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC  
AACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGTCTTCTTTCAGTGTAAGTGTG

3681 DRA3,

3722 ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla  
AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG  
TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTTTCGTCGCCGC

3782 SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis  
TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGACAGCCTGACGCCCCACAC  
AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG

3816 HIND3,

3842 SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla  
TCAGCCAAATCCAAGTTTGTTATGGGGCAAAAGACGTCCGTTGCCATGCCAGAAAGGCC  
AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCGGG

3875 AAT2, 3890 BGLI,

3902 ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp  
GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC  
CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

3962 ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys  
ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTTCAGCCTGAGAAGGGGGGTCGTAAG  
TGATGGTAGTACCGATTCTTGCTCCAAAGACGCAAGTCGGACTCTTCCCCCAGCATTCT

4022 ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu  
CCAGCTCGTCTCATCGTGTTCCTCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG  
GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAC

4082 TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr  
TACGACGTGGTTACAAAGCTCCCTTGCCGTGATGGGAAGCTCCTACGGATTCCAATAC  
ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

4142 SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet  
TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG  
AGTGGTCCTGTGCGCCCACTTAAGGAGCACGTTGCGACCTTCAGGTTCTTTTGGGGTTAC

4160 ECORI,

4202 GlyPh SerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr  
GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG  
CCAAGAGCATACTATGGGCGACGAACTGAGGTGTAGTGAAGTCTCGCTGTAGGCATGC

FIG. 22-7



4229 DRD1, 4236 ALWN1,

4262 GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer  
GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCAAGCCCGCGTGGCCATCAAGTCC  
CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTTCGGGCGCACCGGTAGTTCAGG

4301 BGLI, 4308 BALI,

4322 LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly  
CTACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC  
GAGTGGCTCTCCGAAATACAACCCCGGGGAGAATGGTTAAGTTCCTCCCTCTTGACGCCG

4345 APAI,

4382 TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys  
TATCGCAGGTGCCGCGCGAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTCACTTGC  
ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG

4442 TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal  
TACATCAAGGCCCGGGCAGCCTGTGAGCCGCGAGGGCTCCAGGACTGCACCATGCTCGTG  
ATGTAGTTCCGGGCCCGTCCGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCAC

4452 SMAI XMAI,

4502 CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer  
TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGCGCAGC  
ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCCTGCGCCGCTCG

4508 DRD1, 4511 TTH3I,

4562 LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln  
CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCTGGGGACCCCCACAA  
GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGACCCCTGGGGGGTGT

4622 ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp  
CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCCCACGAC  
GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG

4637 SACI,

4682 GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla  
GGCGCTGGAAAGAGGGTCTACTACCTCACCCTGACCCTACAACCCCTCGCGAGAGCT  
CCGCGACCTTTCTCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA

4731 NRUI,

4742 AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe  
GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT  
CGCACCTCTGTCGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAA

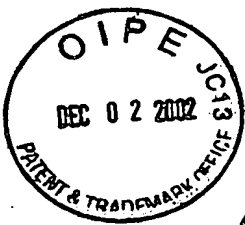
4802 AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla  
GCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC  
CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG

4806 PFLM1, 4807 DRA3,

ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu

FIG. 22-8





4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA  
TCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT  
^

4893 BGL2,

4922 ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis  
CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC  
GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG  
^

4954 NCOI,

4982 SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro  
AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACCTGGGGTACCG  
TCAATGAGAGGTCCACTTTAGTTATCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC  
^ ^

5015 SPHI, 5035 KPNI,

5042 ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly  
CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA  
GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT  
^ ^

5064 APAI, 5091 BALI,

5102 GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys  
GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA  
CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTCGAGTTT  
^

5113 NDEI,

5162 LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr  
CTACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTACGGCTGGCTAC  
GAGTGAGGTTATCGCCGGCGACCGGTGCACCTGAACAGGCCGACCAAGTGCCGACCGATG  
^ ^ ^

5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,

5222 SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys  
AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCGCTGGATCTGGTTTTGC  
TCGCCCCCTCTGTAAATAGTGTGCACAGAGTACGGGCCGGGGCGACCTAGACCAAACG  
^

5240 DRA3,

5282 LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn  
CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCAACCGAATGAGCACGAAT  
GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTTACTCGTGCTTA  
^

5295 PSTI,

5342 ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe  
CCTAAACCTCAAAGAAAGACCAAACGTAACACCAACCGGCCGCGCAGGACGTCAAGTTC  
GGATTTGGAGTTTCTTTCTGTTTGCATTGTGGTTGGCCGCCGCGTCTGCAGTTCAAG  
^ ^ ^

5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,

5402 ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu  
CCGGGTGGCGGTCAGATCGTTGGTGGAGTTTACTTGTGCGCGCAGGGGCCCTAGATTG  
GGCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC  
^

FIG. 22-9



5449 APAI,

5462 GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro  
GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTGCAACCTCGAGGTAGACGTCAGCCT  
CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

5522 IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro  
ATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC  
TAGGGGTTCCGAGCAGCCGGGCTCCCGTCCTGGACCCGAGTCGGGCCCCATGGGAACCGGG

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

5582 LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg  
CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG  
GAGATACCGTTACTCCCGACGCCCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC

5642 ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp  
CCTAGCTGGGGCCCCACAGACCCCCGGCGTAGGTCGCGCAATTTGGGTAAGGTCATCGAT  
GGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCATTCCAGTAGCTA

5650 APAI, 5696 CLAI,

5702 ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValGlyAlaProLeu  
ACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTCGTCGGCGCCCCCTCTT  
TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGCCGCGGGGAGAA

5724 HGIE2, 5750 KAS1 NARI, 5756 ECON1,

5762 GlyGlyAlaAlaArgAlaOC AM  
GGAGGCGCTGCCAGGGCCTAATAGTCGAC  
CCTCCGCGACGGTCCCGGATTATCAGCTG

5785 SALI,

FIG. 22-10